

Local Asymptotic Normality of the spectrum of high-dimensional spiked F-ratios

Prathapasinghe Dharmawansa, Iain M. Johnstone, and Alexei Onatski.

November 17, 2014

Abstract

We consider two types of *spiked* multivariate F distributions: a scaled distribution with the scale matrix equal to a rank-one perturbation of the identity, and a distribution with trivial scale, but rank-one non-centrality. The norm of the rank-one matrix (*spike*) parameterizes the joint distribution of the eigenvalues of the corresponding F matrix. We show that, for a *spike* located above a *phase transition* threshold, the asymptotic behavior of the log ratio of the joint density of the eigenvalues of the F matrix to their joint density under a local deviation from this value depends only on the largest eigenvalue λ_1 . Furthermore, λ_1 is asymptotically normal, and the *statistical experiment* of observing all the eigenvalues of the F matrix converges in the Le Cam sense to a *Gaussian shift experiment* that depends on the asymptotic mean and variance of λ_1 . In particular, the best statistical inference about a sufficiently large *spike* in the local asymptotic regime is based on the largest eigenvalue only. As a by-product of our analysis, we establish joint asymptotic normality of a few of the largest eigenvalues of the multi-spiked F matrix when the corresponding spikes are above the *phase transition* threshold.

KEY WORDS: Spiked F-ratio, Local Asymptotic Normality, multivariate F distribution, phase transition, super-critical regime, asymptotic normality of eigenvalues, limits of statistical experiments.

1 Introduction

In this paper we establish the *Local Asymptotic Normality* (*LAN*) of the *statistical experiments* of observing the eigenvalues of the F-ratio, $B^{-1}A$, of two high-dimensional independent *Wishart* matrices, A and B . We consider two situations. First, both A and B are central Wisharts with dimensionality and degrees of freedom that grow proportionally, and with the covariance parameters that differ by a matrix of rank one. Second, A and B have the same covariance parameter, but A is a non-central Wishart with the non-centrality parameter of rank one. In both cases, the joint distribution of the eigenvalues of $B^{-1}A$ depends on the norm of the rank-one matrix, which we call a *spike*. We

find that the considered *statistical experiments* are *LAN* under a local parameterization of the *spike* when the locality is above a *phase transition* threshold.

Many classical multivariate statistical tests are based on the eigenvalues of F-ratio matrices. For example, all tests of the equality of two covariance matrices and of the *general linear hypothesis* in the *Multivariate Linear Model* described in Muirhead's (1982) chapters 8 and 10 are of this form. Contemporaneous statistical applications often require the dimensionality of the F-ratio and its degrees of freedom be large and comparable. Therefore, we consider the asymptotic regime where the dimensionality and the degrees of freedom diverge to infinity at the same rate.

Our requirement that the parameters of the two Wisharts differ by a rank-one matrix can be linked to situations where the alternative hypothesis is characterized by the presence of one factor or signal, which is absent from the data under the null. Inference conditional on factors requires considering non-central F-ratios, whereas the unconditional inference leads to F-ratios with unequal covariances.

The main result of this paper can be summarized as follows. We show that the asymptotic behavior of the log ratio of the joint density of the eigenvalues of $B^{-1}A$, which corresponds to a sufficiently large value of the *spike*, to their joint density under a local deviation from this value depends only on the largest eigenvalue λ_1 . Furthermore, λ_1 is asymptotically normal, and the *statistical experiment* of observing all the eigenvalues of $B^{-1}A$ converges in the Le Cam sense to a *Gaussian shift experiment* that depends on the asymptotic mean and variance of λ_1 . In particular, the best statistical inference about a sufficiently large *spike* in the local asymptotic regime is based on the largest eigenvalue only.

We derive an explicit formula for the *phase transition* threshold demarcating the area of the sufficiently large *spikes*. In a general framework, where the parameters of A and B may differ by a matrix Δ of a finite rank, we show that, when the norm of Δ is below the threshold, any finite number of the largest eigenvalues of $B^{-1}A$ almost surely converge to the upper boundary of the support of the limiting spectral distribution of $B^{-1}A$, derived by Wachter (1980). In contrast, when m of the largest eigenvalues of Δ are above the threshold, we find that the m of the largest eigenvalues of $B^{-1}A$ almost surely converge to locations strictly above the upper boundary of Wachter's distribution, and that their local fluctuations about these limits are asymptotically jointly normal.

In a setting of two independent and not necessarily normal samples, the *phase transition* phenomenon has been studied in Nadakuditi and Silverstein (2010). They obtain a formula for the threshold, and establish the almost sure limits of the m largest eigenvalues for the case where Δ describes the difference between covariance matrices of the two samples. The limiting distribution of fluctuations above the threshold is described in their paper as an open problem. Our paper solves this problem for the case of two normal samples.

The *phase transition* phenomenon for a single Wishart matrix has also been a subject of active recent research. Baik et al (2005) study the joint distributions of a few of the largest eigenvalues of *complex Wisharts* with spiked covariance parameters. They derive the asymptotic distributions of a few of the largest eigenvalues, which turn out to be different depending on whether the sizes of the corresponding spikes are below, at, or above a *phase transition* threshold, the situations often referred to as the *sub-critical*, *critical*, and *super-critical regimes*.

Similar transition takes place for real Wisharts. Paul (2007) establishes asymptotic normality of the fluctuations of a few of the largest eigenvalues in the *super-critical regime* of the real case. Féral and Pécché (2009), Benaych-Georges et al (2011) and Bao et al (2014a) show that the fluctuations in the *sub-critical* real case have the *Tracy-Widom distribution*, while Mo (2012) and Bloemendal and Viràg (2011, 2013) establish the asymptotic distribution of a different type in the critical regime. In a setting of two normal samples, Bao et al (2014b) study the almost sure limits of the *sample canonical correlations* when the *population canonical correlations* are below and when they are above a *phase transition* threshold.

Our results on the joint asymptotic normality of the largest eigenvalues in the *super-critical regime* for F-ratios can be used to make statistical inference about the eigenvalues of the “ratio” of the population covariances of A and B , or the eigenvalues of the non-centrality parameter of A . The estimates of these eigenvalues play important role in MANOVA and the *discriminant analysis*, and can also be used in constructing modified model selection criteria as discussed in Sheena et al (2004). Further, they may be important in as diverse applications as constructing genetic selection indices and describing a degree of financial turbulence (see Hayes and Hill (1981), and Kritzman and Li (2010)).

We expect that our asymptotic normality results can be extended to the case of the “ratio” of two sample covariance matrices constructed from non-normal samples. In the one-sample case, such an extension of Paul’s (2007) asymptotic normality results has been done in Bai and Yao (2008). In this paper, we focus on normal data. This focus is dictated by our main goal: establishing the *LAN* property of the statistical experiments of observing the eigenvalues of $B^{-1}A$. To reach this goal, we derive an asymptotic approximation to a log likelihood process by representing it in the form of a contour integral, and applying the *Laplace approximation* method. The explicit form of the joint distribution of the eigenvalues of $B^{-1}A$ is known only in the normal case, and we need such an explicit form for our analysis.

A decision-theoretic approach to the finite sample estimation of the eigenvalues of the “ratio” of the population covariances of A and B , or the eigenvalues of the non-centrality parameter of A was taken in many previous studies (see Sheena et al (2004), Bilodeau and Srivastava (1992), and references therein). In one of the first such studies, Muirhead and Verathaworn (1985) explain that the ideal decision-theoretic approach that directly

analyzes expected loss with respect to the joint distribution of the eigenvalues of $B^{-1}A$ “does not seem feasible due primarily to the complexity of the distribution of the ordered latent roots...” Instead, they focus on deriving an optimal estimator from a particular class.

Our *LAN* result makes possible an asymptotic implementation of the ideal decision-theoretic approach. We overcome the complexity of the joint distribution of the eigenvalues by using a tractable contour integral representation of the log likelihood process, which was obtained in the single-spike case by Dharmawansa and Johnstone (2014). In the multiple-spike case, a similar representation involves multiple contour integrals (see Passemier et al (2014)). An asymptotic analysis of such a multiple integral requires a substantial additional effort, and we leave it for future research.

It is interesting to contrast the *LAN* result in the *super-critical* regime with the asymptotic behavior of the log likelihood ratio in the case of a *sub-critical* spike. In a separate research effort, we follow Onatski et al (2013), who analyze the log likelihood ratio in the *sub-critical* regime for the case of a single Wishart matrix, to show that the experiment of observing the eigenvalues of $B^{-1}A$ in the *sub-critical* regime is not of the *LAN* type. Furthermore, the log-likelihood process turns out to depend only on a smooth functional of the empirical distribution of all the eigenvalues of $\Sigma^{-1}A$, so that asymptotically efficient inference procedures may ignore the information contained in λ_1 altogether. The results of this *sub-critical* analysis will be published elsewhere.

The rest of the paper is structured as follows. In the next section, we describe our setting. In Section 3, we derive the almost sure limits of a few of the largest eigenvalues of the F-ratio. In Section 4, we establish the asymptotic normality of the eigenvalue fluctuations in the *super-critical* regime. In Section 5, we derive an asymptotic approximation to the joint distribution of the eigenvalues of $B^{-1}A$ for the special case of a single *super-critical* spike. In Section 6, we show that the likelihood ratio in the local parameter space is asymptotically equivalent to a centered and scaled largest eigenvalue, and establish the *LAN* property. Section 7 concludes.

2 Setup

Suppose that

$$A \sim W_p(n_1 + k, \Sigma_1, \Omega_1) \quad \text{and} \quad B \sim W_p(n_2, \Sigma_2)$$

are independent non-central and central Wishart matrices respectively. For the non-centrality parameter Ω_1 , we use a symmetric version of the definition in Muirhead (1982, p. 442). That is, if Z is an $n \times p$ matrix distributed as $N(M, I_n \otimes \Sigma)$, then $Z'Z \sim W_p(n, \Sigma, \Omega)$ with the non-centrality parameter $\Omega = \Sigma^{-1/2}M'M\Sigma^{-1/2}$. We will consider

two different settings for the parameters Σ_1, Σ_2 , and Ω_1 .

Setting 1 (Spiked covariance) $\Sigma_2 = \Sigma$, $\Sigma_1 = \Sigma^{1/2}(I_p + VhV')\Sigma^{1/2}$, and $\Omega_1 = 0$.

Here $\Sigma^{1/2}$ is the symmetric square root of a positive definite matrix Σ ; V in a $p \times k$ matrix of nuisance parameters with orthonormal columns, and $h = \text{diag}\{h_1, \dots, h_k\}$ is the diagonal matrix of the “covariance spikes” with $h_1 > \dots > h_k$.

Setting 2 (Spiked non-centrality) $\Sigma_2 = \Sigma$, $\Sigma_1 = \Sigma$, and $\Omega_1 = (n_1 + k)VhV'$, where Σ , V , and h are as defined above, but h_j with $j = 1, \dots, k$ are interpreted as “non-centrality spikes.”

We are interested in the behavior of the eigenvalues of

$$\mathbf{F} \equiv (B/n_2)^{-1} A/n_A,$$

where

$$n_A = n_1 + k,$$

as n_1, n_2 , and p grow so that $p/n_1 \rightarrow c_1$ and $p/n_2 \rightarrow c_2$ with $0 < c_i < 1$, while k , the number of spikes, remains fixed. In what follows, we will assume that $\Sigma = I_p$. This assumption is without loss of generality because the eigenvalues of \mathbf{F} do not change under the transformation $A \mapsto \Sigma^{-1/2}A\Sigma^{-1/2}$, $B \mapsto \Sigma^{-1/2}B\Sigma^{-1/2}$.

It is convenient to think of A/n_A as a sample covariance matrix XX'/n_A of the sample X having the factor structure

$$X = V\mathcal{F}' + \varepsilon \tag{1}$$

with V, \mathcal{F} , and ε playing the roles of the factor loadings, factors, and idiosyncratic terms, respectively. Matrices \mathcal{F} and ε are mutually independent, and independent from B . The distribution of ε is $N(0, I_p \otimes I_{n_A})$, and the distribution of \mathcal{F} depends on the setting. For Setting 1, $\mathcal{F} \sim N(0, I_p \otimes h)$, whereas for Setting 2, \mathcal{F} is a deterministic matrix such that $\mathcal{F}'\mathcal{F}/n_A = h$. With this interpretation, Settings 1 and 2 describe, respectively, distributions which are unconditional and conditional on the factors. In both cases the spike parameters h_j , $j = 1, \dots, k$, measure the factors’ variability.

We would like to introduce a convenient representation of the eigenvalues of \mathbf{F} , that we will denote as $\lambda_{p1} \geq \dots \geq \lambda_{pp}$. First, note that λ_{pj} , $j = 1, \dots, p$, are invariant with respect to the simultaneous transformations

$$A \mapsto UAU' \equiv n_A \tilde{H} \quad \text{and} \quad B \mapsto UBU' \equiv n_2 E, \tag{2}$$

where U is a random matrix uniformly distributed over the orthogonal group $\mathcal{O}(p)$. Under the assumption that $\Sigma = I_p$, matrix $n_2 E$ is distributed as $W_p(n_2, I_p)$ and is

independent from \tilde{H} . Matrix \tilde{H} has the form $\tilde{X}\tilde{X}'/n_A$, where

$$\tilde{X} = \tilde{V}\mathcal{F}' + \tilde{\varepsilon}$$

with $\tilde{\varepsilon} \sim N(0, I_p \otimes I_{n_A})$ independent from \tilde{V} , and \tilde{V} being a random $p \times k$ matrix uniformly distributed on the Stiefel manifold of orthogonal k -frames in \mathbb{R}^p . We can think of \tilde{V} as having the form

$$\tilde{V} = v(v'v)^{-1/2} \equiv vW_v^{-1/2},$$

where $v \sim N(0, I_p \otimes I_k)$ and $W_v \equiv v'v$ is Wishart $W_k(p, I_k)$.

Further, let $O_{\mathcal{F}} \in \mathcal{O}(n_A)$ be such that the submatrix of its first k columns equals $\mathcal{F}(\mathcal{F}'\mathcal{F})^{-1/2}$, and let $\hat{X} = \tilde{X}O_{\mathcal{F}}$. Clearly,

$$\tilde{H} = \tilde{X}\tilde{X}'/n_A = \hat{X}\hat{X}'/n_A, \quad (3)$$

and matrix \hat{X} has the form

$$\hat{X} = vW_v^{-1/2}h^{1/2}W_{\mathcal{F}}^{1/2} + \hat{\varepsilon},$$

where $v, W_{\mathcal{F}}$ and $\hat{\varepsilon}$ are mutually independent and independent from E ; $\hat{\varepsilon} \sim N(0, I_p \otimes I_{n_A})$; and the distribution of $W_{\mathcal{F}}$ depends on the setting. For Setting 1, $W_{\mathcal{F}} \sim W_k(n_A, I_k)$, whereas for Setting 2, $W_{\mathcal{F}} = n_A I_k$.

Finally, let us denote the submatrix of the first k columns of $\hat{\varepsilon}$ as u . Then

$$\hat{X}\hat{X}' = \xi\xi' + n_1 H, \quad (4)$$

where $n_1 H \sim W_p(n_1, I_p)$, H and $\xi\xi'$ are mutually independent, and independent from E , and

$$\xi = vW_v^{-1/2}h^{1/2}W_{\mathcal{F}}^{1/2} + u. \quad (5)$$

Using (2), (3), and (4), we obtain the convenient representation for the eigenvalues, announced above. Let $\hat{x}_{p1} \geq \dots \geq \hat{x}_{pp}$ be the roots of the equation

$$\det(\xi\xi'/n_1 + H - xE) = 0. \quad (6)$$

Then

$$\lambda_{pj} = n_1 \hat{x}_{pj} / (n_1 + k). \quad (7)$$

This representation is convenient because the roots of (6) can be viewed and analyzed as perturbations of the roots of equation $\det(H - xE) = 0$ caused by adding the low-rank matrix $\xi\xi'/n_1$ to H .

If $x \in \mathbb{R}$ is such that $H - xE$ is invertible, then

$$(\xi\xi'/n_1 + H - xE)^{-1} = S - S\xi(I_k + \xi'S\xi/n_1)^{-1}\xi'S/n_1,$$

where $S \equiv (H - xE)^{-1}$. Therefore, if x is a root of the equation

$$\det(I_k + \xi'(H - xE)^{-1}\xi/n_1) = 0, \quad (8)$$

then it also solves (6), and hence, the asymptotic behavior of the roots of (6) can be inferred from that of the random matrix-valued function

$$M(x) = \xi'(H - xE)^{-1}\xi/n_1. \quad (9)$$

This is the main idea of the analysis in the next section of the paper.

3 Almost sure limits of the largest eigenvalues

Let $\mathbf{n} \equiv (n_1, n_2)$ and $\mathbf{c} \equiv (c_1, c_2)$. We will denote the asymptotic regime where n_1, n_2 , and p grow so that $p/n_1 \rightarrow c_1$ and $p/n_2 \rightarrow c_2$ with $c_j \in (0, 1)$ as $p, \mathbf{n} \rightarrow_{\mathbf{c}} \infty$. As follows from Wachter's (1980) work, as $p, \mathbf{n} \rightarrow_{\mathbf{c}} \infty$, the empirical distribution of the eigenvalues of $E^{-1}H$ converges in probability to the distribution with density

$$\frac{1 - c_2}{2\pi} \frac{\sqrt{(b_+ - \lambda)(\lambda - b_-)}}{\lambda(c_1 + c_2\lambda)} \mathbf{1}_{\{b_- \leq \lambda \leq b_+\}}. \quad (10)$$

The upper and the lower boundaries of the support of this density are

$$b_{\pm} = \left(\frac{1 \pm r}{1 - c_2} \right)^2, \text{ where } r = \sqrt{c_1 + c_2 - c_1c_2}.$$

The results of Silverstein and Bai (1995) and Silverstein (1995) show that the empirical distribution converges not only in probability, but also almost surely (a.s.). Furthermore, as follows from Theorem 1.1 of Bai and Silverstein (1998), the largest eigenvalue of $E^{-1}H$ a.s. converges to b_+ .

The latter convergence, together with (7) and Weyl's inequalities for the eigenvalues of a sum of two Hermitian matrices (see Theorem 4.3.7 in Horn and Johnson (1985)), imply that the $k + 1$ -th largest eigenvalue of \mathbf{F} , $\lambda_{p,k+1}$, a.s. converges to b_+ . Those of the k largest eigenvalues that remain separated from b_+ as $p, \mathbf{n} \rightarrow_{\mathbf{c}} \infty$, must correspond to solutions of (8). Below, we study these solutions in detail.

Lemma 1 For any $x > b_+$, as $p, \mathbf{n} \rightarrow_{\mathbf{c}} \infty$,

$$\frac{1}{p} \operatorname{tr} \left[(H - xE)^{-1} \right] \xrightarrow{\text{a.s.}} m_x(0) \text{ and} \quad (11)$$

$$\frac{1}{p} \operatorname{tr} \left[\frac{d}{dx} (H - xE)^{-1} \right] \xrightarrow{\text{a.s.}} \frac{d}{dx} m_x(0), \quad (12)$$

where $m_x(0) = \lim_{z \rightarrow 0} m_x(z)$, and $m_x(z) \in \mathbb{C}^+$ is analytic in $z \in \mathbb{C}^+$, and satisfies equation

$$z - \frac{1}{1 + c_1 m_x(z)} = -\frac{1}{m_x(z)} - \frac{x}{1 - c_2 x m_x(z)}. \quad (13)$$

Proof: Let $x \in \mathbb{R}$ be such that $x > b_+$, and let $F_x(\lambda)$ be the empirical distribution function of the eigenvalues of $H - xE$. For any $z \in \mathbb{C}^+$, let

$$\hat{m}_x(z) = \int (\lambda - z)^{-1} dF_x(\lambda)$$

be the Stieltjes transform of $F_x(\lambda)$. Note that matrix $H - xE$ can be represented in the form YTY'/p , where $Y \sim N(0, I_p \otimes I_{n_1+n_2})$ and T is a diagonal matrix with the first n_1 and the last n_2 diagonal elements equal to p/n_1 and $-xp/n_2$, respectively. Therefore, by Theorem 1.1 of Silverstein and Bai (1995), for any $z \in \mathbb{C}^+$, $\hat{m}_x(z)$ a.s. converges to $m_x(z) \in \mathbb{C}^+$, which is an analytic function in the domain $z \in \mathbb{C}^+$ that solves the functional equation (13).

By Theorem 1.1 of Bai and Silverstein (1998), the largest eigenvalue of $E^{-1}H$ a.s. converges to b_+ . Therefore, for any $x > b_+$, the largest eigenvalue of $H - xE$ is a.s. asymptotically bounded away from the positive semi-axis. Hence, $\hat{m}_x(z)$ is analytic and bounded in a small disc D around $z = 0$ for all sufficiently large p and \mathbf{n} , a.s. By Vitali's theorem (see Titchmarsh (1960), p.168), $\hat{m}_x(z)$ is a.s. converging to an analytic function in D . Since, in $D \cap \mathbb{C}^+$, the limiting function is $m_x(z)$, we have

$$\frac{1}{p} \operatorname{tr} \left[(H - xE)^{-1} \right] = \hat{m}_x(0) \xrightarrow{\text{a.s.}} m_x(0),$$

where $m_x(0) = \lim_{z \rightarrow 0} m_x(z)$. Further, $\frac{1}{p} \operatorname{tr} \left[(H - \zeta E)^{-1} \right]$ is an analytic bounded function of ζ in a small disk D_x around x , for all sufficiently large p and \mathbf{n} , a.s. Therefore, by Vitali's theorem its a.s. limit $f(\zeta)$ is analytic in D_x , and

$$\frac{1}{p} \operatorname{tr} \left[\frac{d}{d\zeta} (H - \zeta E)^{-1} \right] \xrightarrow{\text{a.s.}} \frac{d}{d\zeta} f(\zeta)$$

in D_x . On the other hand, we know that $f(\zeta) = m_{\operatorname{Re} \zeta}(0)$ for ζ from D_x . Therefore, we have (12). \square

Lemma 2 For any $x > b_+$, as $p, \mathbf{n} \rightarrow_{\mathbf{c}} \infty$,

$$\left\| M(x) - (h + c_1 I_k) \frac{1}{p} \operatorname{tr} \left[(H - xE)^{-1} \right] \right\| \xrightarrow{a.s.} 0 \text{ and } \left\| \frac{d}{dx} M(x) - (h + c_1 I_k) \frac{1}{p} \operatorname{tr} \left[\frac{d}{dx} (H - xE)^{-1} \right] \right\| \xrightarrow{a.s.} 0,$$

where $\|\cdot\|$ denotes the spectral norm.

Proof: This convergences follow from (5), (9), and Lemma 3 stated below. \square

Lemma 3 Let C be a random $p \times p$ matrix, independent from u and v , which are as defined in Section 2, and such that $p\|C\|$ is bounded for all sufficiently large p , a.s. Then, as $p \rightarrow \infty$,

$$\|v' C v - (\operatorname{tr} C) I_k\| \xrightarrow{a.s.} 0 \text{ and } \|v' C u\| \xrightarrow{a.s.} 0.$$

Proof: This lemma follows from the Borel-Cantelli lemma, and the upper bounds on the fourth moments of the entries $v' C v - (\operatorname{tr} C) I_k$ and $v' C u$ established by Lemma 2.7 of Bai and Silverstein (1998). \square

Lemma 4 (i) For any $\varepsilon > 0$, the k eigenvalues of $M(x)$ are strictly increasing functions of $x \in (b_+ + \varepsilon, \infty)$ for sufficiently large p and \mathbf{n} , a.s.; (ii) $m_x(0)$ is a strictly increasing, continuous function of $x \in (b_+, \infty)$; (iii) $\lim_{x \rightarrow \infty} m_x(0) = 0$, and $\lim_{x \downarrow b_+} m_x(0) (h_i + c_1) < -1$ if and only if $h_i > \bar{h}$, where

$$\bar{h} = \frac{c_2 + r}{1 - c_2}.$$

Proof: Let $\mu_1 \in (0, \infty)$ be the largest eigenvalue of $E^{-1}H$. For any $x_1 > x_2 > \mu_1$, matrix $(x_1 E - H)^{-1} - (x_2 E - H)^{-1}$ is negative definite, a.s. Part (i) follows from this, from the definition (9), and from the fact that $\mu_1 \xrightarrow{a.s.} b_+$. Part (i) together with Lemmas 1 and 2 imply that $m_x(0)$ is increasing on (b_+, ∞) . It is strictly increasing because, otherwise, (13) would not be satisfied for some $z \in \mathbb{C}^+$ that are sufficiently close to zero. The continuity follows from the analyticity of $m_x(0)$ established in the proof of Lemma 1. Finally, $\lim_{x \rightarrow \infty} m_x(0) = 0$ is implied by (ii) and (11). Equation (13) implies that

$$\lim_{x \downarrow b_+} m_x(0) = \frac{c_2 - 1}{(r + 1)r},$$

which, in its turn, implies the second statement of (iii). \square

Let $\hat{x}_{p1} \geq \dots \geq \hat{x}_{pk}$ be the solutions of equation (8). By Lemmas 1, 2, and 4, if

$$h_1 > \dots > h_m > \bar{h} > h_{m+1} > \dots > h_k, \quad (14)$$

then $\hat{x}_{pi} \xrightarrow{a.s.} x_i$, where x_i , $i = 1, \dots, m$, are such that

$$1 + (h_i + c_1) m_{x_i}(0) = 0 \quad (15)$$

and $m_{x_i}(0)$ satisfies (13) with x replaced by x_i . In particular,

$$\frac{1}{1 + c_1 m_{x_i}(0)} - \frac{1}{m_{x_i}(0)} - \frac{x_i}{1 - c_2 x_i m_{x_i}(0)} = 0. \quad (16)$$

Combining (15) and (16), we obtain

$$\frac{1}{h_i} + 1 - \frac{x_i}{h_i + c_1 + c_2 x_i} = 0,$$

which implies that

$$x_i = \frac{(h_i + c_1)(h_i + 1)}{h_i - c_2(h_i + 1)}. \quad (17)$$

By (7), $n_1 \hat{x}_{pi} / (n_1 + k)$, $i = 1, \dots, m$, must be the m largest eigenvalues of \mathbf{F} , and thus, x_i , $i = 1, \dots, m$, describe their a.s. limits. Since there are only m roots of (8) that are asymptotically separated from b_+ and are located above b_+ , the other $k - r$ of the largest eigenvalues of \mathbf{F} must a.s. converge to b_+ . To summarize, the following proposition holds.

Proposition 5 *Suppose that $h_1 > \dots > h_m > \bar{h} > h_{m+1} > \dots > h_k$. Then, for $i \leq m$, the i -th largest eigenvalue of \mathbf{F} a.s. converges to x_i defined in (17). For $m < i \leq k$, the i -th largest eigenvalue a.s. converges to b_+ .*

As follows from Proposition 5, $\bar{h} = (c_2 + r) / (1 - c_2)$ is the phase transition threshold for the eigenvalues of the spiked F-ratio \mathbf{F} . The value of this threshold diverges to infinity when $c_2 \rightarrow 1$. Note that, when c_2 is close to one, the smallest eigenvalue of B/n_2 is close to zero, which makes $(B/n_2)^{-1}$ a particularly bad estimator of the inverse of the population covariance, Σ^{-1} . When $c_2 \rightarrow 0$, the phase transition converges to $\sqrt{c_1}$, which is the phase transition threshold for the eigenvalues of a single spiked Wishart matrix. In such a case, x_i converges to $(h_i + c_1)(h_i + 1) / h_i$, which is the a.s. limit of the i -th largest eigenvalue of the spiked Wishart when the i -th spike h_i is above $\sqrt{c_1}$.

4 Asymptotic normality

In what follows, we will assume that (14) holds, so that only m eigenvalues of \mathbf{F} separate from the bulk asymptotically. We would like to study their fluctuations around the corresponding a.s. limits. Proposition 5 shows that the limits x_i depend on c_1 and c_2 . Because of this dependence, the rate of the convergence has to depend on the rates of the convergences $p/n_1 \rightarrow c_1$ and $p/n_2 \rightarrow c_2$. However, as will be shown below, the latter rates do not affect the fluctuations of λ_{pi} around

$$x_{pi} = \frac{(h_i + c_{p1})(h_i + 1)}{h_i - c_{p2}(h_i + 1)},$$

which are obtained from x_i by replacing c_1 and c_2 by $c_{p1} = p/n_1$ and $c_{p2} = p/n_2$.

Similar to x_i , which are linked to the Stieltjes transform of the limiting spectral distribution of $xE - H$ via (15), x_{pi} also can be linked to the limiting Stieltjes transform, albeit under a slightly different asymptotic regime. Precisely, let $m_{px}(z)$ be the Stieltjes transform of the limiting spectral distribution of $xE - H$ as n_1, n_2 , and p grow so that p/n_1 and p/n_2 remain fixed. Then, similarly to (15), we have

$$1 + (h_i + c_{p1}) m_{px_{pi}}(0) = 0. \quad (18)$$

This link will be useful in our analysis below, where we maintain the assumption that p/n_1 and p/n_2 are not necessarily fixed, but converge to c_1 and c_2 , respectively.

Recall that, by (7), $\lambda_{pi} = n_1 \hat{x}_{pi} / (n_1 + k)$, where $\hat{x}_{pi}, i = 1, \dots, m$, satisfy (8). Clearly, the asymptotic distributions of $\sqrt{p}(\lambda_{pi} - x_{pi})$ and $\sqrt{p}(\hat{x}_{pi} - x_{pi})$, $i = 1, \dots, m$, coincide. Therefore, below we will study the asymptotic behavior of $\sqrt{p}(\hat{x}_{pi} - x_{pi})$, $i = 1, \dots, m$. By the standard Taylor expansion argument,

$$\sqrt{p}(\hat{x}_{pi} - x_{pi}) = - \frac{\sqrt{p} \det \mathcal{M}(x_{pi})}{\frac{d}{dx} \det \mathcal{M}(x_{pi}) + \frac{1}{2}(\hat{x}_{pi} - x_{pi}) \frac{d^2}{dx^2} \det \mathcal{M}(\tilde{x}_{pi})}, \quad (19)$$

$i = 1, \dots, m$, where $\mathcal{M}(x) = I_k + M(x)$, and $\tilde{x}_{pi} \in [x_{pi}, \hat{x}_{pi}]$. We have

$$\frac{d}{dx} \det \mathcal{M}(x_{pi}) = \det \mathcal{M}(x_{pi}) \operatorname{tr} S(x_{pi}),$$

and

$$\frac{d^2}{dx^2} \det \mathcal{M}(x_{pi}) = \det \mathcal{M}(x_{pi}) \left\{ \operatorname{tr} R(x_{pi}) + (\operatorname{tr} S(x_{pi}))^2 - \operatorname{tr} [S^2(x_{pi})] \right\},$$

where

$$S(x) = \mathcal{M}(x)^{-1} \frac{d}{dx} M(x), \text{ and } R(x) = \mathcal{M}(x)^{-1} \frac{d^2}{dx^2} M(x).$$

Since the event

$$\det \mathcal{M}(x_{pi}) = 0 \text{ or } 1 + M_{ii}(x_{pi}) = 0 \text{ for some } i = 1, \dots, m$$

happens with probability zero, we can simultaneously multiply the numerator and denominator of (19) by $(1 + M_{ii}(x_{pi})) / \det \mathcal{M}(x_{pi})$ to obtain

$$\sqrt{p}(\hat{x}_{pi} - x_{pi}) = - \frac{\sqrt{p}(1 + M_{ii}(x_{pi}))}{s(x_{pi}) + \frac{1}{2}(\hat{x}_{pi} - x_{pi}) \delta(x_{pi})}, \quad (20)$$

where

$$s(x_{pi}) = (1 + M_{ii}(x_{pi})) \operatorname{tr} S(x_{pi}),$$

and

$$\delta(x_{pi}) = (1 + M_{ii}(x_{pi})) \left\{ \text{tr } R(x_{pi}) + (\text{tr } S(x_{pi}))^2 - \text{tr } [S^2(x_{pi})] \right\}.$$

Lemma 6 *For any $i = 1, \dots, m$, we have: (i) $s(x_{pi}) \xrightarrow{P} (h_i + c_1) \frac{d}{dx} m_{x_i}(0)$; (ii) $\delta(x_{pi}) = O(1)$ a.s.*

Proof: By Lemmas 1 and 2,

$$\frac{d}{dx} M(x_{pi}) \xrightarrow{a.s.} (h + c_1 I_k) \frac{d}{dx} m_{x_i}(0). \quad (21)$$

Further,

$$(1 + M_{ii}(x_{pi})) (I_k + M(x_{pi}))^{-1} \xrightarrow{a.s.} \text{diag } \{0, \dots, 0, 1, 0, \dots, 0\} \quad (22)$$

with 1 at the i -th place on the diagonal. The latter convergence follows from the fact that $I_k + M(x_{pi})$ can be viewed as a small perturbation of a diagonal matrix

$$I_k + (h + c_1 I_k) m_{x_i}(0),$$

which has non-zero diagonal elements, except at the i -th position. The eigenvalue perturbation formulae (see, for example, (2.33) on p.79 of Kato (1980)) will then lead to (22). Combining (21) and (22), and using the definition of $s(x_{pi})$, we obtain (i).

To establish (ii), we note that $(1 + M_{ii}(x_{pi})) \text{tr } R(x_{pi}) = O_P(1)$ by an argument similar to that used to establish (i). Further, $(\text{tr } S(x_{pi}))^2 - \text{tr } [S^2(x_{pi})]$ is a linear function of the only eigenvalue of $S(x_{pi})$ that diverges to infinity. By the eigenvalue perturbation formulae, such an eigenvalue equals $(1 + M_{ii}(x_{pi}))^{-1} O(1)$ a.s. Therefore,

$$(1 + M_{ii}(x_{pi})) \left((\text{tr } S(x_{pi}))^2 - \text{tr } [S^2(x_{pi})] \right) = O(1),$$

which concludes the proof of (ii). \square

Equation (20), Lemma 6, and the Slutsky theorem imply that, for the purpose of establishing convergence in distribution of $\sqrt{p}(\hat{x}_{pi} - x_{pi})$, $i = 1, \dots, m$, we may focus on the numerator of (20)

$$Z_{ii}(x_{pi}) \equiv \sqrt{p}(1 + M_{ii}(x_{pi})) = \sqrt{p}(M_{ii}(x_{pi}) - (h_i + c_{p1}) m_{px_{pi}}(0)),$$

where the last equality follows from (18).

The random variable Z_{ii} is the entry of the matrix

$$Z(x_{pi}) = \sqrt{p}(M(x_{pi}) - (h + c_{p1} I_k) m_{px_{pi}}(0))$$

that belongs to the i -th row and the i -th column. Let us now introduce new notations.

Let

$$\begin{aligned} D &= (W_{\mathcal{F}}/n_1)^{1/2} h^{1/2} (W_v/p)^{-1/2}, \\ G &= (H - x_{pi}E)^{-1} / p, \\ \Delta_{\mathcal{F}} &= \sqrt{n_1} \left((W_{\mathcal{F}}/n_1)^{1/2} - I_k \right), \text{ and} \\ \Delta_v &= \sqrt{p} (W_v/p - I_k). \end{aligned}$$

Then, using equations (9) and (5), we obtain the following decomposition.

$$Z(x_{pi}) = \sum_{v=1}^6 Z^{(v)},$$

where

$$\begin{aligned} Z^{(1)} &= D \sqrt{p} (v' G v - I_k \operatorname{tr} G) D', \\ Z^{(2)} &= (\operatorname{tr} G) D (W_v/p)^{-1/2} h^{1/2} \sqrt{c_{p1}} \Delta_{\mathcal{F}}, \\ Z^{(3)} &= \operatorname{tr} G \sqrt{c_{p1}} \Delta_{\mathcal{F}} h^{1/2} (W_v/p)^{-1} h^{1/2}, \\ Z^{(4)} &= -(\operatorname{tr} G) h^{1/2} \Delta_v (W_v/p)^{-1} h^{1/2}, \\ Z^{(5)} &= \sqrt{c_{p1}} \sqrt{p} (D v' G u + u' G v D'), \\ Z^{(6)} &= c_{p1} \sqrt{p} (u' G u - I_k \operatorname{tr} G), \end{aligned}$$

and

$$Z^{(7)} = (h + c_{p1} I_k) \sqrt{p} (\operatorname{tr} G - m_{px_{pi}}(0)).$$

For the last term, $Z^{(7)}$, we prove the following lemma.

Lemma 7 $Z^{(7)} \xrightarrow{P} 0$.

Proof: The proof of this lemma will appear in a separate work. Had x_{pi} been negative, $H - x_{pi}E$ would have been having the form YTY' with $Y \sim N(0, I_p \otimes I_{n_1+n_2})$ and a positive definite diagonal T with converging spectral distribution. The Lemma would have been following then from the results of Bai and Silverstein (2004). Our proof extends Bai and Silverstein's (2004) arguments to the case of negative x_{pi} . \square

Further, the asymptotic behavior of the terms $Z^{(2)}$ and $Z^{(3)}$ differ depending on the setting. Recall that for Setting 1, $W_{\mathcal{F}} \sim W_k(n_A, I_k)$. Then, since

$$\Delta_{\mathcal{F}} = \sqrt{n_1} (W_{\mathcal{F}}/n_1 - I_k) / 2 + o_P(1),$$

a standard CLT together with Lemma 1 imply that

$$Z^{(2)} + Z^{(3)} \xrightarrow{d} N(0, 2c_1 m_{x_i}^2(0) h^2). \quad (23)$$

The latter limit is independent from the limits of $Z^{(j)}$, $j \neq 2, 3$, because $W_{\mathcal{F}}$ is independent from u and v .

In contrast, for Setting 2, we have $W_{\mathcal{F}} = n_A I_k$, and $\Delta_{\mathcal{F}} = o(1)$. Therefore,

$$Z^{(2)} + Z^{(3)} \xrightarrow{P} 0. \quad (24)$$

Let us now establish the convergence of $Z^{(j)}$, $j \leq 6$ such that $j \neq 2, 3$. Let l_i and L_i be such that $[l_i, L_i]$ includes the support of the limiting spectral distribution, G_{x_i} , of $H - x_{pi}E$. Moreover, let $[l_i, L_i]$ be such that none of the eigenvalues $\lambda_{p1}^{(i)}, \dots, \lambda_{pp}^{(i)}$ of $H - x_{pi}E$ lies outside $[l_i, L_i]$ for sufficiently large p , a.s. Further, let g_q with $q = 1, \dots, Q$, where Q is an arbitrary positive integer, be functions which are continuous on $[l_i, L_i]$ and let ζ denote a $p \times m$ matrix with i.i.d. $N(0, 1)$ entries. Finally, let

$$\Theta = \{(q, s, t) : q = 1, \dots, Q; 1 \leq s \leq t \leq m\}.$$

The following Lemma is a slight modification of Lemma 13 of the Supplementary Appendix in Onatski (2012).

Lemma 8 *The joint distribution of random variables*

$$\left\{ \frac{1}{\sqrt{p}} \sum_{j=1}^p g_q \left(\lambda_{pj}^{(i)} \right) (\zeta_{js} \zeta_{jt} - \delta_{st}), (q, s, t) \in \Theta \right\}$$

weakly converges to a multivariate normal. The covariance between components (q, s, t) and (q_1, s_1, t_1) of the limiting distribution is equal to 0 when $(s, t) \neq (s_1, t_1)$, and to $(1 + \delta_{st}) \int g_q(\lambda) g_{q_1}(\lambda) dG_{x_i}(\lambda)$ when $(s, t) = (s_1, t_1)$.

Proof: For readers' convenience, we provide a proof of this Lemma in the Appendix. \square

Note that all entries of $Z^{(j)}$, $j \leq 6$ such that $j \neq 2, 3$, are linear combinations of the terms having the form considered in Lemma 8, with weights converging in probability to finite constants. Take, for example $Z^{(1)}$. Its entries are linear combinations of the entries of

$$\frac{1}{\sqrt{p}} v' (H - x_{pi}E)^{-1} v - I_k \frac{1}{\sqrt{p}} \text{tr} (H - x_{pi}E)^{-1},$$

which, in turn, can be represented in the form $\frac{1}{\sqrt{p}} \sum_{j=1}^p \left(\lambda_{pj}^{(i)} \right)^{-1} (\zeta_{js} \zeta_{jt} - \delta_{st})$. The matrix ζ is obtained by multiplying $[u, v]$ from the left by the eigenvector matrix of $H - x_{pi}E$.

Lemma 8 implies that vector $(Z_{ii}^{(1)}, Z_{ii}^{(4)}, Z_{ii}^{(5)}, Z_{ii}^{(6)})$ converges in distribution to a

four-dimensional normal vector with zero mean and the following covariance matrix

$$\begin{pmatrix} 2h_i^2 m'_{x_i}(0) & -2h_i^2 m_{x_i}^2(0) & 0 & 0 \\ -2h_i^2 m_{x_i}^2(0) & 2h_i^2 m_{x_i}^2(0) & 0 & 0 \\ 0 & 0 & 4c_1 h_i m'_{x_i}(0) & 0 \\ 0 & 0 & 0 & 2c_1^2 m'_{x_i}(0) \end{pmatrix}.$$

Combining this result with Lemma 7, and convergencies (23), and (24), we obtain, for Setting 1,

$$Z_{ii}(x_{pi}) \xrightarrow{d} N\left(0, 2(h_i + c_1)^2 m'_{x_i}(0) - 2h_i^2 (1 - c_1) m_{x_i}^2(0)\right), \quad (25)$$

and, for Setting 2,

$$Z_{ii}(x_{pi}) \xrightarrow{d} N\left(0, 2(h_i + c_1)^2 m'_{x_i}(0) - 2h_i^2 m_{x_i}^2(0)\right). \quad (26)$$

To establish the joint convergence of $Z_{ii}(x_{pi})$, $i = 1, \dots, m$, we need another lemma. For each $i = 1, \dots, m$, let $g_q^{(i)}$, with $q = 1, \dots, Q$, be functions continuous on $[l_i, L_i]$.

Lemma 9 *For any set of pairs $\{(s_i, t_i) : i = 1, \dots, m\}$ such that $(s_{i_1}, t_{i_1}) \neq (s_{i_2}, t_{i_2})$ for any $i_1 \neq i_2$, the joint distribution of random variables*

$$\left\{ \frac{1}{\sqrt{p}} \sum_{j=1}^p g_q^{(i)}(\lambda_{pj}^{(i)}) (\zeta_{js_i} \zeta_{jt_i} - \delta_{s_i t_i}), i = 1, \dots, m \right\}$$

weakly converges to a multivariate normal. The covariance between components i_1 and i_2 of the limiting distribution is equal to 0 when $i_1 \neq i_2$.

The proof of this lemma is very similar to that of Lemma 8, and we omit it to save space. Lemma 9 implies that $Z_{ii}(x_{pi})$, $i = 1, \dots, m$ jointly converge to an m -dimensional normal vector with a diagonal covariance matrix. This result, together with equation (20), Lemma 6, and convergences (25, 26) establish the following Lemma.

Lemma 10 *The joint asymptotic distribution of $\sqrt{p}(\lambda_{pi} - x_{pi})$, $i = 1, \dots, m$ is normal, with diagonal covariance matrix. For Setting 1, the i -th diagonal element of the covariance matrix equals*

$$\frac{2(h_i + c_1)^2 m'_{x_i}(0) - 2h_i^2 (1 - c_1) m_{x_i}^2(0)}{(h_i + c_1)^2 \left(\frac{d}{dx} m_{x_i}(0)\right)^2}. \quad (27)$$

For Setting 2, it equals

$$\frac{2(h_i + c_1)^2 m'_{x_i}(0) - 2h_i^2 m_{x_i}^2(0)}{(h_i + c_1)^2 \left(\frac{d}{dx} m_{x_i}(0)\right)^2}. \quad (28)$$

In the Appendix, we establish the following explicit expressions for $m_{x_i}^2(0)$, $m'_{x_i}(0)$, and $\frac{d}{dx}m_{x_i}(0)$:

$$m_{x_i}^2(0) = (h_i + c_1)^{-2}, \quad (29)$$

$$m'_{x_i}(0) = -\frac{h_i^2}{(h_i + c_1)^2 \left(c_1 + c_2(1 + h_i)^2 - h_i^2 \right)}, \quad (30)$$

$$dm_{x_i}(0)/dx = \frac{-(c_2(1 + h_i) - h_i)^2}{(h_i + c_1)^2 \left(c_1 + c_2(1 + h_i)^2 - h_i^2 \right)}. \quad (31)$$

Using (29), (30), and (31) in (27) and (28), we obtain

Proposition 11 *For any $h_1 > \dots > h_m > \bar{h} \equiv (c_2 + r)/(1 - c_2)$, the joint asymptotic distribution of $\sqrt{p}(\lambda_{pi} - x_{pi})$, $i = 1, \dots, m$ is normal with diagonal covariance matrix. For Setting 1,*

$$\sqrt{p}(\lambda_{pi} - x_{pi}) \xrightarrow{d} N \left(0, 2r^2 \frac{h_i^2 (h_i + 1)^2 \left(h_i^2 - c_2(h_i + 1)^2 - c_1 \right)}{(c_2 - h_i + c_2 h_i)^4} \right), \quad (32)$$

whereas for Setting 2,

$$\sqrt{p}(\lambda_{pi} - x_{pi}) \xrightarrow{d} N \left(0, 2t^2 \frac{h_i^2 (h_i + 1)^2 \left(h_i^2 - c_2(h_i + 1)^2 - c_1 \right)}{(c_2 - h_i + c_2 h_i)^4} \right). \quad (33)$$

Here

$$r^2 = c_1 + c_2 - c_1 c_2,$$

$$t^2 = c_1 + c_2 - \frac{c_1 (h_i^2 - c_1)}{(1 + h_i)^2},$$

and

$$x_{pi} = \frac{(h_i + p/n_1)(h_i + 1)}{h_i - (h_i + 1)p/n_2}.$$

Remark 12 *It is straightforward to verify that $t^2 < r^2$ as long as $h_i > \bar{h}$. Therefore, the asymptotic variance of λ_i is smaller for Setting 2 than for Setting 1. This accords with intuition because, as discussed above, Setting 2 corresponds to the asymptotic analysis conditional on factors \mathcal{F} , whereas Setting 1 corresponds to the unconditional asymptotic analysis. The factors' variance adds to the asymptotic variance of λ_i .*

Remark 13 *For Setting 1, when $c_2 \rightarrow 0$, the asymptotic variance of λ_i converges to the correct asymptotic variance*

$$2c_1 (h_i + 1)^2 (h_i^2 - c_1) / h_i^2$$

of the largest eigenvalue of the spiked Wishart model. Non-centrality spikes in Wishart distribution were considered in Onatski (2007). The limit of the asymptotic variance in (33) when $c_2 \rightarrow 0$ coincides with the formula for the asymptotic variance derived there.

5 Analysis of the joint density of eigenvalues

From now on, let us consider the case of a single spike, which is located above the phase transition threshold \bar{h} . That is, assume that $k = m = 1$, and let $h_1 = h_p$. We would like to study the asymptotic behavior of the ratio of the joint densities of all the eigenvalues of \mathbf{F} that correspond to

$$H_0 : h_p = h_0 \text{ and to } H_1 : h_p = h_0 + \gamma/\sqrt{p},$$

where $h_0 > \bar{h}$ is fixed and γ is a local parameter.

Following James (1964) and Khatri (1967), we can write the joint density of the eigenvalues of \mathbf{F} in Setting 1 as

$$f_1(\Lambda; h_p) = \frac{Z_{p1}(\Lambda)}{(1 + h_p)^{n_A/2}} {}_1F_0 \left(\frac{n}{2}; \frac{h_p}{h_p + 1} VV', \alpha_p \Lambda (I_p + \alpha_p \Lambda)^{-1} \right),$$

and in Setting 2 as

$$f_2(\Lambda; h) = \frac{Z_{p2}(\Lambda)}{\exp \{h_p n_A/2\}} {}_1F_1 \left(\frac{n}{2}; \frac{n_A}{2}; \frac{n_A h_p}{2} VV', \alpha_p \Lambda (I_p + \alpha_p \Lambda)^{-1} \right),$$

where ${}_1F_0$ and ${}_1F_1$ are the hypergeometric functions of two matrix arguments, $\alpha_p = n_A/n_2$, $n = n_A + n_2$, $\Lambda = \text{diag} \{\lambda_{p1}, \dots, \lambda_{pp}\}$, and $Z_{pj}(\Lambda)$, $j = 1, 2$, depend on n_A, n_2, p and Λ , but not on h_p . The joint densities are evaluated at the observed values of the eigenvalues.

To facilitate analysis, we use Proposition 1 of Dharmawansa and Johnstone (2014) to rewrite $f_1(\Lambda; h_p)$ and $f_2(\Lambda; h_p)$ as shown in the following lemma.

Lemma 14 *Consider the region $\mathbb{C} \setminus (1, \infty)$ in the complex plane. Let $\tilde{\mathcal{K}}$ be a contour defined in that region which starts at $-\infty$, encircles $\tilde{\lambda}_{pj} = \alpha_p \lambda_{pj} / (1 + \alpha_p \lambda_{pj})$, $j = 1, \dots, p$, counter-clockwise and returns to $-\infty$. Then we have*

$$f_1(\Lambda; h_p) = C_{p1}(\Lambda) k_{p1}(h_p) \frac{1}{2\pi i} \int_{\tilde{\mathcal{K}}} \left(1 - \frac{h_p}{1 + h_p} z \right)^{\frac{p-n-2}{2}} \prod_{j=1}^p (z - \tilde{\lambda}_{pj})^{-\frac{1}{2}} dz, \quad (34)$$

and

$$f_2(\Lambda; h_p) = C_{p2}(\Lambda) k_{p2}(h_p) \frac{1}{2\pi i} \times \int_{\tilde{\mathcal{K}}} {}_1F_1 \left(\frac{n-p+2}{2}, \frac{n_A-p+2}{2}; \frac{n_A h_p}{2} z \right) \prod_{j=1}^p (z - \tilde{\lambda}_{pj})^{-\frac{1}{2}} dz \quad (35)$$

where $C_{pj}(\Lambda)$, $j = 1, 2$, depend on n_A, n_2, p and Λ , but not on h_p , $k_{p1}(h_p) = (1 + h_p)^{\frac{p-2-n_A}{2}} h_p^{1-p/2}$, and $k_{p2}(h_p) = \exp \{-n_A h_p/2\} h_p^{1-p/2}$.

We will now derive an asymptotic approximation to the contour integrals in (34) and (35). First, we will analyze (34) and then turn to (35).

5.1 Asymptotic approximation: Setting 1

Let us deform the contour $\tilde{\mathcal{K}}$, without changing the integral's value with probability approaching one as $p, \mathbf{n} \rightarrow_{\mathbf{c}} \infty$, as shown in Figure 1. Let $\mathcal{K} = \mathcal{K}^+ \cup \overline{\mathcal{K}^+}$ with $\mathcal{K}^+ = \mathcal{K}_1^+ \cup \mathcal{K}_2^+ \cup \mathcal{K}_3^+ \cup \mathcal{K}_4^+$, where

$$\begin{aligned} \mathcal{K}_1^+ &= \{z : \Im(z) \geq 0 \text{ and } |z - \tilde{\lambda}_{p1}| = \epsilon\}, \\ \mathcal{K}_2^+ &= \{z : z \in [\tilde{x}_0, \tilde{\lambda}_{p1} - \epsilon]\}, \\ \mathcal{K}_3^+ &= \{z : \Re(z) = \tilde{x}_0 \text{ and } 0 \leq \Im(z) \leq \tilde{x}_0\}, \text{ and} \\ \mathcal{K}_4^+ &= \{z : \Re(z) \leq \tilde{x}_0 \text{ and } \Im(z) = \tilde{x}_0\}. \end{aligned}$$

Here $\epsilon > 0$ is a small number and $\alpha b_+ / (1 + \alpha b_+) < \tilde{x}_0 < \alpha x_1 / (1 + \alpha x_1)$ with

$$\alpha = \lim \alpha_p = c_2 / c_1,$$

and

$$x_1 = \lim x_{p1} = \frac{(h_0 + c_1)(h_0 + 1)}{h_0 - (h_0 + 1)c_2}. \quad (36)$$

As follows from our results in the previous section, $\tilde{\lambda}_{p1} \xrightarrow{a.s.} \alpha x_1 / (1 + \alpha x_1)$ and $\tilde{\lambda}_{p2} \xrightarrow{a.s.} \alpha b_+ / (1 + \alpha b_+)$, so $\tilde{x}_0 \in (\tilde{\lambda}_{p2}, \tilde{\lambda}_{p1})$ for sufficiently large p and \mathbf{n} , a.s.

Consider the following integral over the deformed contour \mathcal{K}

$$I_p(\gamma, \Lambda) = \int_{\mathcal{K}} {}_1\mathcal{F}_0(z) \prod_{j=1}^p (z - \tilde{\lambda}_{pj})^{-\frac{1}{2}} dz, \quad (37)$$

where

$${}_1\mathcal{F}_0(z) \equiv \left(1 - \frac{h_p}{1 + h_p} z\right)^{\frac{p-n-2}{2}}.$$

For two sequences of random variables $\{\xi_p\}$ and $\{\eta_p\}$, we will write $\xi_p \stackrel{P}{\sim} \eta_p$ if and only

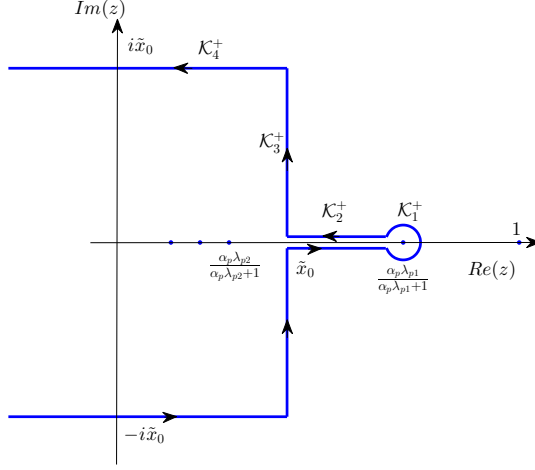


Figure 1: The contour \mathcal{K} .

if ξ_p/η_p converges in probability to 1 as $p, \mathbf{n} \rightarrow_{\mathbf{c}} \infty$. We have the following lemma.

Lemma 15 *Under the hypothesis that $h_p = h_0$, uniformly in γ from any compact subset of \mathbb{R}*

$$I_p(\gamma, \Lambda) \stackrel{\mathbb{P}}{\sim} {}_1\mathcal{F}_0 \left(\tilde{\lambda}_{p1} \right) 2 \left(-\frac{\pi}{pH_0} \right)^{\frac{1}{2}} \prod_{j=2}^p \left(\tilde{\lambda}_{p1} - \tilde{\lambda}_{pj} \right)^{-\frac{1}{2}},$$

where the principal branches of the square roots are used, and

$$H_0 = \frac{h_0(1 - c_2)(\mu_0 - \sqrt{b_+})(\mu_0 - \sqrt{b_-})(c_1 + c_2\mu_0)}{2c_1c_2(h_0 - c_2\mu_0)\mu_0(c_1 + h_0)} \quad (38)$$

with $\mu_0 = h_0 + 1$.

Proof: Let $\mathcal{K}_j = \mathcal{K}_j^+ \cup \overline{\mathcal{K}_j^+}$ for $j = 1, \dots, 4$. Using this notation, we can decompose (37) as

$$I_p(\gamma, \Lambda) = I_{1,2,p}(\gamma, \Lambda) + I_{3,4,p}(\gamma, \Lambda) \quad (39)$$

where $I_{1,2,p}(\gamma, \Lambda)$ is the part of the integral corresponding to $\mathcal{K}_1 \cup \mathcal{K}_2$, and $I_{3,4,p}(\gamma, \Lambda)$ is the part corresponding to the rest of the contour, $\mathcal{K}_3 \cup \mathcal{K}_4$. Our strategy is to show that the integral $I_p(\gamma, \Lambda)$ is asymptotically equivalent to $I_{1,2,p}(\gamma, \Lambda)$, the integral $I_{3,4,p}(\gamma, \Lambda)$ being asymptotically dominated by $I_{1,2,p}(\gamma, \Lambda)$.

Let us first focus on $I_{1,2,p}(\gamma, \Lambda)$. Since the singularity of the integrand at $\tilde{\lambda}_{p1}$ is of the inverse square root type, as the radius ϵ of \mathcal{K}_1 converges to zero, the integral over

\mathcal{K}_1 converges to zero too. Therefore, we have

$$I_{1,2,p}(\gamma, \Lambda) = 2I_{2,p}(\gamma, \Lambda), \quad (40)$$

where

$$I_{2,p}(\gamma, \Lambda) = \int_{\tilde{\lambda}_{p1}}^{\tilde{x}_0} {}_1\mathcal{F}_0(z) \prod_{j=1}^p (z - \tilde{\lambda}_{pj})^{-\frac{1}{2}} dz.$$

Changing the variable of integration from z to $x = \tilde{\lambda}_{p1} - z$, we arrive at

$$\begin{aligned} I_{2,p}(\theta, \lambda) &= - \int_0^{\tilde{\lambda}_{p1} - \tilde{x}_0} {}_1\mathcal{F}_0(\tilde{\lambda}_{p1} - x) \prod_{j=1}^p (\tilde{\lambda}_{p1} - \tilde{\lambda}_{pj} - x)^{-\frac{1}{2}} dx \\ &= {}_1\mathcal{F}_0(\tilde{\lambda}_{p1}) (-1)^{\frac{1}{2}} \int_0^{\tilde{\lambda}_{p1} - \tilde{x}_0} \exp\{-pf_p(x)\} x^{-\frac{1}{2}} dx \end{aligned}$$

where

$$f_p(x) = \frac{\beta_p}{2} \ln \left(1 + \frac{h_p(1 + \alpha_p \lambda_{p1})}{1 + h_p + \alpha_p \lambda_{p1}} x \right) + \frac{1}{2p} \sum_{j=2}^p \ln (\tilde{\lambda}_{p1} - \tilde{\lambda}_{pj} - x)$$

with $\beta_p = (2 + n - p)/p$. Now the integral $\int_0^{\tilde{\lambda}_{p1} - \tilde{x}_0} \exp\{-pf_p(x)\} x^{-\frac{1}{2}} dx$ can be evaluated using standard Laplace approximation steps (see Olver (1997), section 7.3) as follows.

First, let us show that the derivative $\frac{d}{dx} f_p(x)$ is continuous and positive on $x \in [0, \tilde{\lambda}_{p1} - \tilde{x}_0]$ for sufficiently large p and \mathbf{n} , a.s. We have

$$\frac{d}{dx} f_p(x) = \frac{\beta_p}{2} \frac{h_p(1 + \alpha_p \lambda_{p1})}{(h_p + (1 + \alpha_p \lambda_{p1})(1 + h_p x))} - \frac{1}{2p} \sum_{j=2}^p (\tilde{\lambda}_{p1} - \tilde{\lambda}_{pj} - x)^{-1}.$$

Therefore, the continuity follows from the fact that, when $h_p = h_0$,

$$\lambda_{p1} \xrightarrow{a.s.} x_1 \equiv \frac{(h_0 + c_1)(h_0 + 1)}{h_0 - (h_0 + 1)c_2} \text{ and } \lambda_{p2} \xrightarrow{a.s.} b_+.$$

In order to establish the positivity, we first obtain

$$\min_{x \in [0, \tilde{\lambda}_{p1} - \tilde{x}_0]} \frac{d}{dx} f_p(x) = \frac{\beta_p}{2} \frac{h_p}{1 + h_p - h_p \tilde{x}_0} - \frac{1}{2p} \sum_{j=2}^p (\tilde{x}_0 - \tilde{\lambda}_{pj})^{-1}.$$

It is straightforward to verify that the above equation can be represented in the following

form

$$\begin{aligned} \min_{x \in [0, \tilde{\lambda}_{p1} - \tilde{x}_0]} \frac{d}{dx} f_p(x) &= \frac{\beta_p}{2} \frac{h_p(1 + \alpha x_0)}{1 + h_p + \alpha x_0} + \frac{p-1}{2p} (1 + \alpha x_0) \\ &\quad + \frac{1}{2\alpha_p p} (1 + \alpha x_0)^2 \sum_{j=2}^p (\lambda_{pj} - \alpha x_0 / \alpha_p)^{-1}. \end{aligned}$$

where $x_0 = \tilde{x}_0 / (\alpha(1 - \tilde{x}_0))$. Therefore, we obtain

$$\min_{x \in [0, \tilde{\lambda}_{p1} - \tilde{x}_0]} \frac{d}{dx} f_p(x) \xrightarrow{a.s.} \Psi(x_0, h_0) \quad (41)$$

where

$$\begin{aligned} \Psi(x_0, h_0) &= \frac{\beta h_0(1 + \alpha x_0)}{2(1 + h_0 + \alpha x_0)} + \frac{1}{2} (1 + \alpha x_0) + \frac{1}{2\alpha} (1 + \alpha x_0)^2 m(x_0), \\ \beta &= c_1^{-1} + c_2^{-1} - 1, \end{aligned}$$

and $m(x_0) = \lim_{z \rightarrow x_0} m(z)$ with $m(z)$ being the Stieltjes transform of the limiting spectral distribution of \mathbf{F} , that is the distribution with density (10).

Since $m(x_0)$ is increasing on $x_0 \in (b_+, \infty)$, we have

$$\Psi(x_0, h_0) > \lim_{x_0 \downarrow b_+} \Psi(x_0, h_0).$$

Moreover, noting the fact that $\lim_{x_0 \downarrow b_+} \Psi(x_0, h_0)$ is an increasing function of h_0 and $h_0 > \bar{h} \equiv (c_2 + r) / (1 - c_2) = \sqrt{b_+} - 1$, we obtain

$$\Psi(x_0, h_0) > \lim_{x_0 \downarrow b_+} \Psi(x_0, h_0) > \lim_{x_0 \downarrow b_+} \Psi(x_0, \bar{h}). \quad (42)$$

Finally, direct calculations, which are not reported here to save space, show that, as x_0 converges to b_+ from the right,

$$m(x_0) \rightarrow -1/(b_+ - \sqrt{b_+}). \quad (43)$$

This in turn gives

$$\lim_{x \downarrow b_+} \Psi(x, \bar{h}) = 0 \quad (44)$$

which establishes the positivity.

Since $\lambda_{p1} \xrightarrow{a.s.} x_1$, we have

$$f'_p(0) \equiv \left. \frac{d}{dx} f_p(x) \right|_{x=0} \xrightarrow{a.s.} H_0,$$

where

$$H_0 = \frac{r^2}{2c_1c_2} \frac{(1 + \alpha x_1)h_0}{1 + h_0 + \alpha x_1} + \frac{1}{2}(1 + \alpha x_1) + \frac{1}{2\alpha}(1 + \alpha x_1)^2 m(x_1).$$

Direct calculations show that

$$m(x_1) = \lim_{z \rightarrow x_1} m(z) = -(1 + h_0)/(x_1 h_0), \quad (45)$$

which, after some algebraic manipulations, gives (38).

We may now exploit the approach given in Olver (1997, pp. 81-82) to yield

$$\int_0^{\tilde{\lambda}_{p1} - \tilde{x}_0} e^{-pf_p(x)} x^{-\frac{1}{2}} dx \stackrel{P}{\sim} \left(\frac{\pi}{pf'_p(0)} \right)^{\frac{1}{2}} e^{-pf_p(0)}.$$

Therefore, we obtain

$$I_{2,p}(\gamma, \Lambda) \stackrel{P}{\sim} {}_1\mathcal{F}_0(\tilde{\lambda}_{p1}) \left(-\frac{\pi}{pH_0} \right)^{\frac{1}{2}} \prod_{j=2}^p (\tilde{\lambda}_{p1} - \tilde{\lambda}_{pj})^{-\frac{1}{2}}. \quad (46)$$

As Lemma 16 below shows, $I_{3,4,p}(\gamma, \Lambda)$ is asymptotically dominated by $I_{2,p}(\theta, \lambda)$, which completes the proof. \square

Lemma 16 *Under the hypothesis that $h_p = h_0$, uniformly in γ from any compact subset of \mathbb{R}*

$$I_{3,4,p}(\gamma, \Lambda) = o_P(I_{2,p}(\gamma, \Lambda)). \quad (47)$$

Proof: Let us first consider the integral over the contour \mathcal{K}_3 . For $z \in \mathcal{K}_3$, we have

$$\left| {}_1\mathcal{F}_0(z) \prod_{j=1}^p (z - \tilde{\lambda}_{pj})^{-\frac{1}{2}} \right| < {}_1\mathcal{F}_0(\tilde{\lambda}_{p1}) e^{-pf_p(\tilde{\lambda}_{p1} - \tilde{x}_0)} (\tilde{\lambda}_{p1} - \tilde{x}_0)^{-\frac{1}{2}}.$$

Also, in view of (41), (42), and (44), we have $f_p(\tilde{\lambda}_{1p} - \tilde{x}_0) > f_p(0) + \epsilon$, for sufficiently large p and \mathbf{n} , a.s., where $\epsilon > 0$. Therefore, using (46), we conclude

$$\int_{\mathcal{K}_3} {}_1\mathcal{F}_0(z) \prod_{j=1}^p (z - \tilde{\lambda}_{pj})^{-\frac{1}{2}} dz = o_P(I_{2,p}(\gamma, \Lambda)). \quad (48)$$

Now consider the integral over the contour \mathcal{K}_4 . We have

$$\begin{aligned} \left| \int_{\mathcal{K}_4} {}_1\mathcal{F}_0(z) \prod_{j=1}^p (z - \tilde{\lambda}_{pj})^{-\frac{1}{2}} dz \right| &< 2 \prod_{j=1}^p |\tilde{\lambda}_{pj} - \tilde{x}_0|^{-\frac{1}{2}} \int_{-\infty}^{\tilde{x}_0} {}_1\mathcal{F}_0(x) dx \\ &= 4 \frac{1 + h_p}{h_p(n - p)} \left(1 - \frac{h_p}{1 + h_p} \tilde{x}_0 \right) {}_1\mathcal{F}_0(\tilde{x}_0) \prod_{j=1}^p |\tilde{\lambda}_{pj} - \tilde{x}_0|^{-\frac{1}{2}}. \end{aligned}$$

Since

$${}_1\mathcal{F}_0(\tilde{x}_0) \prod_{j=1}^p \left(\tilde{\lambda}_{pj} - \tilde{x}_0 \right)^{-\frac{1}{2}} = {}_1\mathcal{F}_0(\tilde{\lambda}_{p1}) e^{-pf_p(\tilde{\lambda}_{p1} - \tilde{x}_0)} \left(\tilde{\lambda}_{p1} - \tilde{x}_0 \right)^{-\frac{1}{2}},$$

we can follow a similar procedure to that outlined above to obtain

$$\int_{\mathcal{K}_4} {}_1\mathcal{F}_0(z) \prod_{j=1}^p \left(z - \tilde{\lambda}_{pj} \right)^{-\frac{1}{2}} dz = o_P(I_{2,p}(\gamma, \Lambda)).$$

This along with (48) gives (47). \square

5.2 Asymptotic approximation: Setting 2

Consider the following integral

$$J_p(\gamma, \Lambda) = \int_{\tilde{\mathcal{K}}} {}_1\mathcal{F}_1(\zeta) \prod_{j=1}^p \left(z - \tilde{\lambda}_{pj} \right)^{-\frac{1}{2}} dz,$$

where

$${}_1\mathcal{F}_1(\zeta) \equiv {}_1F_1(n_A u + 1, n_A v + 1; n_A \zeta)$$

with

$$u = \frac{n_A + n_2 - p}{2n_A}, \quad v = \frac{n_A - p}{2n_A}, \quad \text{and } \zeta = \frac{h_p}{2} z.$$

In Johnstone and Onatski (2014) (Theorem 5), the following result is derived. As $p, \mathbf{n} \rightarrow_{\mathbf{c}} \infty$,

$${}_1\mathcal{F}_1(\zeta) = \frac{C(p, \mathbf{n}) e^{-n_A \varphi(\zeta)}}{\sqrt{2\pi n_A} i} \left(\psi(\zeta) + O(n_A^{-1}) \right), \quad (49)$$

where $O(n_A^{-1})$ is uniform for ζ that do not approach zero or negative semi-axis and $\Re(\zeta) \geq -2u + v$,

$$C(p, \mathbf{n}) = \frac{\Gamma(n_A v + 1) \Gamma(n_A(u - v) + 1)}{\Gamma(n_A u + 1)},$$

$$\varphi(\zeta) = (u - v) \ln(u - v) + v \ln(z_+ + v) - u \ln(z_+ + u) - z_+, \quad (50)$$

where the principal branches of the logarithms are chosen,

$$z_+ = \frac{1}{2} \left\{ \zeta - v + \sqrt{(\zeta - v)^2 + 4u\zeta} \right\}, \quad (51)$$

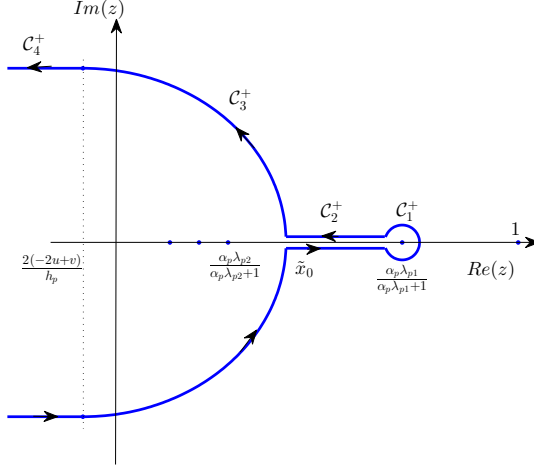


Figure 2: The contour \mathcal{C} .

where the principal branch of the square root is chosen when $\Re(\zeta) \geq -2u + v$ and the other branch is chosen when $\Re \zeta < -2u + v$, and

$$\psi(\zeta) = \left[(z_+ - \zeta) \sqrt{\frac{u}{z_+^2} - \frac{u-v}{(z_+ - \zeta)^2}} \right]^{-1},$$

where the branch of the square root is chosen so that $\sqrt{-1} = -i$.

We will deform the contour $\tilde{\mathcal{K}}$, without changing the integral's value with probability approaching one as $p, \mathbf{n} \rightarrow_{\mathbf{c}} \infty$, as shown in Figure 2. Formally, $\mathcal{C} = \mathcal{C}^+ \cup \overline{\mathcal{C}^+}$ with $\mathcal{C}^+ = \mathcal{C}_1^+ \cup \mathcal{C}_2^+ \cup \mathcal{C}_3^+ \cup \mathcal{C}_4^+$, where

$$\begin{aligned} \mathcal{C}_1^+ &= \{z : \Im(z) \geq 0 \text{ and } |z - \tilde{\lambda}_{p1}| = \epsilon\}, \\ \mathcal{C}_2^+ &= \{z : z \in [\tilde{x}_0, \tilde{\lambda}_{p1} - \epsilon]\}, \\ \mathcal{C}_3^+ &= \{z : z = 2\zeta/h_p \text{ s.t. } \Re(\zeta) \geq -2u + v, \Im(\zeta) \geq 0, |z_+ + u| = |z_{0+} + u|\}, \\ \mathcal{C}_4^+ &= \{z : z = 2\zeta/h_p \text{ s.t. } \Re(\zeta) < -2u + v, \text{ and } \Im(z_+) = |z_{0+} + u|\}. \end{aligned}$$

Here

$$\begin{aligned} z_{0+} &= \frac{1}{2} \left\{ \zeta_0 - v + \sqrt{(\zeta_0 - v)^2 + 4u\zeta_0} \right\}, \text{ and} \\ \zeta_0 &= \frac{h_p}{2} \tilde{x}_0. \end{aligned} \tag{52}$$

Lemma 17 *Under the hypothesis that $h_p = h_0$, uniformly in γ from any compact subset of \mathbb{R}*

$$J_p(\gamma, \Lambda) \stackrel{P}{\sim} 2Z(p, \mathbf{n}, h_0) e^{-n_A \varphi\left(\frac{h_p}{2} \tilde{\lambda}_{p1}\right)} \left(-\frac{\pi}{pH_0}\right)^{1/2} \prod_{j=2}^p \left(\tilde{\lambda}_{p1} - \tilde{\lambda}_{pj}\right)^{-\frac{1}{2}},$$

where

$$Z(p, \mathbf{n}, h_0) = \frac{C(p, \mathbf{n})}{\sqrt{\pi p}} \frac{c_1 + c_2 \mu_0}{\sqrt{c_2 \mu_0^2 + c_1^2 - c_1 + 2c_1 \mu_0}},$$

and H_0 and μ_0 are as defined in Lemma 15.

Proof: Similar to the case of Setting 1, we split $J_p(\gamma, \Lambda)$ into two parts

$$J_p(\gamma, \Lambda) = J_{1,2,p}(\gamma, \Lambda) + J_{3,4,p}(\gamma, \Lambda),$$

where $J_{1,2,p}(\gamma, \Lambda)$ is the part of the integral corresponding to $\mathcal{C}_1 \cup \mathcal{C}_2$, and $J_{3,4,p}(\theta, \lambda)$ is the part corresponding to the rest of the contour, $\mathcal{C}_3 \cup \mathcal{C}_4$. Furthermore,

$$J_{1,2,p}(\gamma, \Lambda) \stackrel{P}{\sim} 2J_{2,p}(\gamma, \Lambda), \quad (53)$$

where

$$J_{2,p}(\gamma, \Lambda) = \int_{\tilde{\lambda}_{p1}}^{\tilde{x}_0} \frac{C(p, \mathbf{n}) e^{-n_A \varphi(\zeta)}}{\sqrt{2\pi n_A i}} \psi(\zeta) \prod_{j=1}^p \left(z - \tilde{\lambda}_{pj}\right)^{-\frac{1}{2}} dz.$$

In contrast to (40), we only have the asymptotic equivalence in (53) because we are using the uniform asymptotic approximation (49) to define $J_{2,p}(\gamma, \Lambda)$.

After the change of the variable of integration, $\zeta \mapsto x = \frac{h_p}{2} \tilde{\lambda}_{p1} - \zeta$, we obtain

$$J_{2,p}(\gamma, \Lambda) = - \int_0^{\frac{h_p}{2}(\tilde{\lambda}_{p1} - \tilde{x}_0)} \frac{C(p, \mathbf{n}) e^{-n_A \varphi(\zeta)}}{\sqrt{2\pi n_A i}} \psi(\zeta) \frac{2}{h_p} \prod_{j=1}^p \left(\frac{2}{h_p} \zeta - \tilde{\lambda}_{pj}\right)^{-\frac{1}{2}} dx.$$

This can be rewritten as

$$J_{2,p}(\gamma, \Lambda) = \left(-\frac{2}{h_p}\right)^{\frac{1}{2}} \frac{C(p, \mathbf{n})}{\sqrt{2\pi n_A i}} \int_0^{\frac{h_p}{2}(\tilde{\lambda}_{p1} - \tilde{x}_0)} e^{-n_A g_p(\zeta)} \psi(\zeta) x^{-\frac{1}{2}} dx,$$

where

$$g_p(\zeta) = \varphi(\zeta) + \frac{1}{2n_A} \sum_{j=2}^p \ln \left(\frac{2}{h_p} \zeta - \tilde{\lambda}_{pj}\right),$$

and

$$\zeta = \frac{h_p}{2} \tilde{\lambda}_{p1} - x.$$

Following the approach in the above analysis in the case of Setting 1, we now would like to show that the derivative $\frac{d}{dx} g_p(\frac{h_p}{2} \tilde{\lambda}_{p1} - x)$ is continuous and positive on $x \in [0, \frac{h_p}{2}(\tilde{\lambda}_{p1} - \tilde{x}_0)]$ for sufficiently large p and \mathbf{n} , a.s. This is equivalent to showing that

$\frac{d}{d\zeta}g_p(\zeta)$ is continuous and *negative* on $\zeta \in [\frac{h_p}{2}\tilde{x}_0, \frac{h_p}{2}\tilde{\lambda}_{p1}]$ for sufficiently large p and \mathbf{n} , a.s.

It is straightforward to verify that z_+ satisfies the quadratic equation

$$z_+^2 + (v - \zeta)z_+ - u\zeta = 0,$$

and

$$\zeta = \frac{z_+(z_+ + v)}{z_+ + u}. \quad (54)$$

From this, and the definition (51) of z_+ , we obtain that, $z_+ > \zeta$ for positive ζ , and

$$\frac{d}{d\zeta}z_+ = \frac{z_+ + u}{2z_+ + v - \zeta} = \frac{(u + z_+)^2}{uv + 2uz_+ + z_+^2} > 0. \quad (55)$$

On the other hand,

$$\frac{d}{dz_+}\varphi(\zeta) = \frac{v}{z_+ + v} - \frac{u}{z_+ + u} - 1 = -\frac{uv + 2uz_+ + z_+^2}{(v + z_+)(u + z_+)} < 0$$

Thus, $\varphi(\zeta)$ is strictly decreasing function of ζ . Furthermore, it is a convex function of $\zeta > 0$. Indeed,

$$\frac{d^2}{dz_+^2}\varphi(\zeta) = -\frac{v}{(z_+ + v)^2} + \frac{u}{(z_+ + u)^2} = \frac{(u - v)(z_+^2 - uv)}{(z_+ + v)^2(z_+ + u)^2},$$

and, using (54) and (55), we also have

$$\frac{d^2}{d\zeta^2}z_+ = -2u(u + z_+)^3 \frac{u - v}{(uv + 2uz_+ + z_+^2)^3}.$$

Therefore, we obtain

$$\begin{aligned} \frac{d^2}{d\zeta^2}\varphi(\zeta) &= \frac{d^2}{dz_+^2}\varphi(\zeta) \left(\frac{d}{d\zeta}z_+ \right)^2 + \frac{d}{dz_+}\varphi(\zeta) \frac{d^2}{d\zeta^2}z_+ \\ &= \frac{(u + z_+)^2(u - v)}{(v + z_+)^2(uv + 2uz_+ + z_+^2)} > 0. \end{aligned}$$

Therefore, $\varphi(\zeta)$ is, indeed, convex for positive ζ , and has a continuous derivative.

Further, it is straightforward to see that

$$w(\zeta) \equiv \frac{1}{2n_A} \sum_{j=2}^p \ln \left(\frac{2}{h_p} \zeta - \tilde{\lambda}_{pj} \right)$$

is a strictly increasing concave function of $\zeta > \frac{h_p}{2} \tilde{\lambda}_{p2}$. This implies that

$$\begin{aligned} \max_{\zeta \in [\frac{h_p}{2} \tilde{x}_0, \frac{h_p}{2} \tilde{\lambda}_{p1}]} \frac{d}{d\zeta} g_p(\zeta) &< \frac{d}{d\zeta} \varphi(\zeta) \Big|_{\zeta=\frac{h_p}{2} \tilde{\lambda}_{1p}} + \frac{d}{d\zeta} w(\zeta) \Big|_{\zeta=\frac{h_p}{2} \tilde{x}_0} \\ &= -\frac{uv + 2uz_+ + z_+^2}{(v + z_+)(u + z_+)} \left(1 + \frac{u(u-v)}{2z_+(u + z_+)}\right) \Big|_{\zeta=\frac{h_p}{2} \tilde{\lambda}_{1p}} \\ &\quad - \frac{2p}{h_p n_A} \frac{p-1}{2p} (1 + \alpha x_0) \\ &\quad - \frac{2p}{h_p n_A} \frac{1}{2\alpha_p p} (1 + \alpha x_0)^2 \sum_{j=2}^p (\lambda_j - \alpha x_0 / \alpha_p)^{-1} \end{aligned}$$

The right hand side of the latter equality a.s. converges to

$$\Pi(x_0, h_0) = -\frac{c_1}{c_2(h_0 + 1)} - 1 - \frac{c_1}{h_0} (1 + \alpha x_0) - \frac{c_1}{h_0 \alpha} (1 + \alpha x_0)^2 m(x_0),$$

where $m(z)$ is the Stieltjes transform of the limiting spectral distribution of \mathbf{F} . Since $m(x_0)$ is an increasing function of $x_0 > b_+$,

$$\Pi(x_0, h_0) < \lim_{x_0 \downarrow b_+} \Pi(x_0, h_0).$$

On the other hand, using (43), we get

$$\lim_{x_0 \downarrow b_+} \Pi(x_0, h_0) = -\frac{c_1}{c_2(h_0 + 1)} - 1 + r \frac{(r + c_2)^2}{c_2 h_0 (1 - c_2)(r + 1)}. \quad (56)$$

Note that, considered as a function of $h_0 > \bar{h}$, $\lim_{x_0 \downarrow b_+} \Pi(x_0, h_0)$ may have positive derivative only when $\lim_{x_0 \downarrow b_+} \Pi(x_0, h_0) < 0$. Indeed,

$$\begin{aligned} \frac{d}{dh_0} \lim_{x_0 \downarrow b_+} \Pi(x_0, h_0) &= \frac{c_1}{c_2(h_0 + 1)^2} - r \frac{(r + c_2)^2}{c_2 h_0^2 (1 - c_2)(r + 1)} \\ &< \frac{1}{h_0} \left(\frac{c_1}{c_2(h_0 + 1)} - r \frac{(r + c_2)^2}{c_2 h_0 (1 - c_2)(r + 1)} \right) \end{aligned}$$

If the latter expression is positive for $h_0 > \bar{h} > 0$, then $\lim_{x_0 \downarrow b_+} \Pi(x_0, h_0)$ is clearly negative. Therefore,

$$\lim_{x_0 \downarrow b_+} \Pi(x_0, h_0) < \max \left\{ 0, \lim_{x_0 \downarrow b_+} \Pi(x_0, \bar{h}) \right\}.$$

But, using the definition $\bar{h} = (c_2 + r) / (1 - c_2)$ in (56), we obtain

$$\lim_{x_0 \downarrow b_+} \Pi(x_0, \bar{h}) = -\frac{c_1(1 - c_2)}{c_2(1 + r)} - 1 + r \frac{(r + c_2)}{c_2(r + 1)} = 0.$$

This implies that $\max_{\zeta \in [\frac{h_p}{2}\tilde{x}_0, \frac{h_p}{2}\tilde{\lambda}_{1p}]} \frac{d}{d\zeta} g_p(\zeta)$ is a.s. negative for sufficiently large p and \mathbf{n} .

Now, since $\lambda_{p1} \xrightarrow{a.s.} x_1$, we have

$$g_p'(0) \equiv -\frac{d}{d\zeta} g_p(\zeta) \Big|_{\zeta = \frac{h_p}{2}\tilde{\lambda}_{p1}} \xrightarrow{a.s.} R_0,$$

where

$$R_0 = \frac{c_1}{c_2(h_0 + 1)} + 1 + \frac{c_1}{h_0}(1 + \alpha x_1) + \frac{c_1}{\alpha h_0}(1 + \alpha x_1)^2 m(x_1).$$

Using (45), (36) and (38), we obtain

$$R_0 = 2c_1 H_0 / h_0.$$

Exploiting the approach given in Olver (1997, pp. 81-82), we obtain

$$\int_0^{\frac{h_p}{2}(\tilde{\lambda}_{p1} - \tilde{x}_0)} e^{-n_A g_p(\zeta)} \psi(\zeta) x^{-\frac{1}{2}} dx \stackrel{\mathbb{P}}{\sim} \left(\frac{\pi}{n_A g_p'(0)} \right)^{\frac{1}{2}} e^{-n_A g_p(\frac{h_p}{2}\tilde{\lambda}_{p1})} \psi\left(\frac{h_0 \alpha x_1}{2(1 + \alpha x_1)} \right).$$

On the other hand, direct calculation shows that

$$\psi\left(\frac{h_0 \alpha x_1}{2(1 + \alpha x_1)} \right) = i \frac{\sqrt{2}(c_1 + c_2 + c_2 h_0)}{\sqrt{c_1} \sqrt{c_2 \mu_0^2 + c_1^2 - c_1 + 2c_1 \mu_0}},$$

and

$$g_p\left(\frac{h_p}{2}\tilde{\lambda}_{p1} \right) = \varphi\left(\frac{h_p}{2}\tilde{\lambda}_{p1} \right) + \frac{1}{2n_A} \sum_{j=2}^p \ln(\tilde{\lambda}_{p1} - \tilde{\lambda}_{pj}).$$

Therefore,

$$J_{2,p}(\gamma, \Lambda) \stackrel{\mathbb{P}}{\sim} Z(p, \mathbf{n}, h_0) \left(-\frac{\pi}{pH_0} \right)^{\frac{1}{2}} e^{-n_A \varphi(\frac{h_p}{2}\tilde{\lambda}_{p1})} \prod_{j=2}^p (\tilde{\lambda}_{p1} - \tilde{\lambda}_{pj})^{-\frac{1}{2}}, \quad (57)$$

where

$$Z(p, \mathbf{n}, h_0) = \frac{C(p, \mathbf{n})}{\sqrt{\pi p}} \frac{c_1 + c_2 \mu_0}{\sqrt{c_2 \mu_0^2 + c_1^2 - c_1 + 2c_1 \mu_0}}.$$

As Lemma 18 below shows, $J_{3,4,p}(\gamma, \Lambda)$ is asymptotically dominated by $J_{2,p}(\theta, \lambda)$, which completes the proof. \square

Lemma 18 *Under the hypothesis that $h_p = h_0$, uniformly in γ from any compact subset of \mathbb{R}*

$$J_{3,4,p}(\gamma, \Lambda) = o_{\mathbb{P}}(J_{2,p}(\gamma, \Lambda)).$$

Proof: Let us first consider the integral $J_{3,p}(\gamma, \Lambda)$ over the contour \mathcal{C}_3 . For $z \in \mathcal{C}_3$, by definition, we have $\Re(\zeta) \equiv \Re(h_p z / 2) \geq -2u + v$. Therefore, the uniform approximation

(49) is still valid, and we have

$$\begin{aligned} J_{3,p}(\gamma, \Lambda) &\stackrel{\text{P}}{\sim} \int_{\mathcal{C}_3} \frac{C(p, \mathbf{n}) e^{-n_A \varphi(\zeta)}}{\sqrt{2\pi n_A i}} \psi(\zeta) \prod_{j=1}^p (z - \tilde{\lambda}_{pj})^{-\frac{1}{2}} dz \\ &= \int_{\mathcal{C}_3} \frac{C(p, \mathbf{n}) e^{-n_A g_p(\zeta)}}{\sqrt{2\pi n_A i}} \psi(\zeta) (z - \tilde{\lambda}_{p1})^{-\frac{1}{2}} dz. \end{aligned} \quad (58)$$

Let us show that, for $\zeta = h_p z/2$ with $z \in \mathcal{C}_3$,

$$\Re g_p(\zeta) > g_p(h_p \tilde{x}_0/2). \quad (59)$$

Recall that

$$\begin{aligned} g_p(\zeta) &= \varphi(\zeta) + \frac{1}{2n_A} \sum_{j=2}^p \ln \left(\frac{2}{h_p} \zeta - \tilde{\lambda}_{pj} \right), \text{ and} \\ \varphi(\zeta) &= (u - v) \ln(u - v) + v \ln(z_+ + v) - u \ln(z_+ + u) - z_+. \end{aligned}$$

By definition of \mathcal{C}_3 , as z moves along \mathcal{C}_3 away from \tilde{x}_0 , ζ is changing so that z_+ moves along a circle with center at $-u$ and radius $z_{0+} + u$, where z_{0+} is as defined in (52). In particular, $|z_+ + u|$ remains constant, $\Re(-z_+)$ increases, and, since $v < u$, $|z_+ + v|$ increases too. Overall,

$$\Re(\varphi(\zeta)) = (u - v) \ln(u - v) + v \ln|z_+ + v| - u \ln|z_+ + u| + \Re(-z_+)$$

is increasing. Note also that $|\zeta| = |z_+||z_+ + v|/|z_+ + u|$ must increase, which implies that $\left| \frac{2}{h_p} \zeta - \tilde{\lambda}_{pj} \right|$ is increasing for all $j \geq 2$, and thus,

$$\Re \left(\frac{1}{2n_A} \sum_{j=2}^p \ln \left(\frac{2}{h_p} \zeta - \tilde{\lambda}_{pj} \right) \right)$$

is increasing too. This implies (59).

On the other hand, in the above proof of Lemma 17 we have shown that $\frac{d}{d\zeta} g_p(\zeta)$ is continuous and *negative* on $\zeta \in [\frac{h_p}{2} \tilde{x}_0, \frac{h_p}{2} \tilde{\lambda}_{p1}]$. Hence, there must exist $C > 0$ such that, for any $\zeta = h_p z/2$ with $z \in \mathcal{C}_3$,

$$\left| e^{-n_A g_p(\zeta)} \right| \leq e^{-n_A C} e^{-n_A \varphi\left(\frac{h_p}{2} \tilde{\lambda}_{p1}\right)} \prod_{j=2}^p \left(\tilde{\lambda}_{p1} - \tilde{\lambda}_{pj} \right)^{-\frac{1}{2}}.$$

This inequality, together with (57) and (58) imply that

$$J_{3,p}(\gamma, \Lambda) = o_P(J_{2,p}(\gamma, \Lambda)).$$

The fact that $J_{4,p}(\gamma, \Lambda) = o_P(J_{2,p}(\gamma, \Lambda))$ follows from pp. 29-31 of Johnstone and Onatski (2014). \square

6 Local Asymptotic Normality

6.1 Analysis for Setting 1

Let us denote the likelihood ratio by

$$L_{p1}(\gamma, \Lambda) = \frac{f_1(\Lambda; h_p)}{f_1(\Lambda; h_0)}. \quad (60)$$

From Lemmas 14 and 15, we obtain the following expression

$$f_1(\Lambda; h_p) = \frac{1}{2\pi i} C_{p1}(\Lambda) k_{p1}(h_p) I_p(\gamma, \Lambda).$$

Using Lemma 15, we obtain

$$L_{p1}(\gamma, \Lambda) \stackrel{P}{\sim} \frac{k_{p1}(h_p)}{k_{p1}(h_0)} \left(\frac{1 - \frac{h_p}{1+h_p} \tilde{\lambda}_{p1}}{1 - \frac{h_0}{1+h_0} \tilde{\lambda}_{p1}} \right)^{\frac{p-n-2}{2}}. \quad (61)$$

Consider a new local parameter

$$\theta_1 = \gamma / \omega_1(h_0),$$

where

$$\omega_1(h_0) = \frac{2h_0^2(1+h_0)^2 r^2}{(h_0 - c_2(1+h_0))^2}.$$

We have the following lemma.

Lemma 19 *Let Under the null hypothesis that $h = h_0$, uniformly in θ_1 from any compact subset of \mathbb{R} ,*

$$\ln L_{p1}(\gamma, \Lambda) = \theta_1 \sqrt{p}(\lambda_{p1} - x_{p1}) - \frac{1}{2} \theta_1^2 \tau_1^2(h_0) + o_P(1)$$

where

$$x_{p1} = \frac{(h_0 + p/n_1)(h_0 + 1)}{h_0 - (h_0 + 1)p/n_2}, \text{ and}$$

$$\tau_1^2(h_0) = 2r^2 \frac{h_0^2(h_0 + 1)^2 (h_0^2 - c_2(h_0 + 1)^2 - c_1)}{(c_2 - h_0 + c_2 h_0)^4}.$$

Proof: Taking the logarithm of (61) yields

$$\begin{aligned} \ln L_{p1}(\gamma, \Lambda) &= \frac{n+2-p}{2} \left(\ln \left(1 - \frac{\tilde{\lambda}_{p1} h_0}{1+h_0} \right) - \ln \left(1 - \frac{\tilde{\lambda}_{p1} h_p}{1+h_p} \right) \right) \\ &\quad - \left(\frac{p-2}{2} \right) \ln \frac{h_p}{h_0} + \left(\frac{p-n_1-2}{2} \right) \ln \frac{1+h_p}{1+h_0} + o_P(1). \end{aligned} \quad (62)$$

Moreover, we have the following expansions

$$\begin{aligned} \ln \left(1 - \frac{\tilde{\lambda}_{p1} h_0}{1+h_0} \right) - \ln \left(1 - \frac{\tilde{\lambda}_{p1} h_p}{1+h_p} \right) &= p^{-\frac{1}{2}} \gamma \frac{\tilde{\lambda}_{p1}}{(1+h_0)(1+h_0(1-\tilde{\lambda}_{p1}))} \\ &\quad - p^{-1} \gamma^2 \frac{\tilde{\lambda}_{p1}}{(1+h_0)^2(1+h_0(1-\tilde{\lambda}_{p1}))} \\ &\quad + p^{-1} \gamma^2 \frac{\tilde{\lambda}_{p1}^2}{2(1+h_0)^2(1+h_0(1-\tilde{\lambda}_{p1}))^2} + o_P(p^{-1}), \end{aligned} \quad (63)$$

$$\ln \frac{1+h_p}{1+h_0} = p^{-\frac{1}{2}} \gamma \frac{1}{1+h_0} - p^{-1} \gamma^2 \frac{1}{2(1+h_0)^2} + o(p^{-1}), \quad (64)$$

and

$$\ln \frac{h_p}{h_0} = p^{-\frac{1}{2}} \gamma h_0^{-1} - \frac{1}{2} p^{-1} \gamma^2 h_0^{-2} + o(p^{-1}). \quad (65)$$

Finally, using (63), (64), and (65) in (62) and noting the fact that $\lambda_1 - x_{p1} \xrightarrow{a.s.} 0$, we obtain the statement of the lemma by straightforward algebraic manipulations. \square

Lemma 19 together with the asymptotic normality of $\sqrt{p}(\lambda_{p1} - x_{p1})$ established in Proposition 11 imply, via Le Cam's First Lemma (see van der Vaart (1998), p.88), that the sequences of the probability measures $\{\mathbb{P}_{h_0,p}\}$ and $\{\mathbb{P}_{h_0+\gamma/\sqrt{p},p}\}$ describing the joint distribution of the eigenvalues of \mathbf{F} under the null $H_0 : h_p = h_0$ and under the local alternative $H_1 : h_p = h_0 + \gamma/\sqrt{p}$ are mutually contiguous. Moreover, the experiments $(\mathbb{P}_{h_0+\theta_1\omega_1(h_0)/\sqrt{p},p} : \theta_1 \in \mathbb{R})$ converge to the Gaussian shift experiment $(N(\theta_1, \tau_1^2(h_0)) : \theta_1 \in \mathbb{R})$. In particular, these experiments are LAN.

6.2 Analysis for Setting 2

Let us denote the likelihood ratio by

$$L_{p2}(\gamma, \Lambda) = \frac{f_2(\Lambda; h_p)}{f_2(\Lambda; h_0)}. \quad (66)$$

From Lemmas 14 and 17, we obtain the following expression

$$f_2(\Lambda; h_p) = \frac{1}{2\pi i} C_{p2}(\Lambda) k_{p2}(h_p) J_p(\gamma, \Lambda).$$

Using Lemma 17, and the definitions (50) and (51), we obtain

$$L_{p2}(\gamma, \Lambda) \stackrel{\text{P}}{\sim} \exp \left[-n_A \sum_{j=1}^4 (a_j(h_p) - a_j(h_0)) \right], \quad (67)$$

where

$$\begin{aligned} a_1(h) &= \frac{h + \ln h}{2}, \\ a_2(h) &= -\frac{1}{2} \left(\frac{h}{2} \tilde{\lambda}_{p1} - v + \sqrt{\left(\frac{h}{2} \tilde{\lambda}_{p1} - v \right)^2 + 4u \frac{h}{2} \tilde{\lambda}_{p1}} \right), \\ a_3(h) &= -u \ln \left[\frac{1}{2} \left(\frac{h}{2} \tilde{\lambda}_{p1} - v + \sqrt{\left(\frac{h}{2} \tilde{\lambda}_{p1} - v \right)^2 + 4u \frac{h}{2} \tilde{\lambda}_{p1}} \right) \right], \end{aligned}$$

and

$$a_4(h) = (u - v) \ln \left[\frac{1}{2} \left(-\frac{h}{2} \tilde{\lambda}_{p1} - v + \sqrt{\left(\frac{h}{2} \tilde{\lambda}_{p1} - v \right)^2 + 4u \frac{h}{2} \tilde{\lambda}_{p1}} \right) \right].$$

We would like, first, to expand $a_j(h_p) - a_j(h_0)$, with $j = 1, \dots, 4$, in the power series of γ/\sqrt{p} up to, and including, the terms of order $O_P\left(\frac{1}{p}\right)$. For a_1 , we have

$$a_1(h_p) - a_1(h_0) = \frac{h_0 + 1}{2h_0} \frac{\gamma}{\sqrt{p}} - \frac{1}{4h_0^2} \frac{\gamma^2}{p}. \quad (68)$$

For a_2 , note that

$$\begin{aligned} \left(\frac{h_p}{2} \tilde{\lambda}_{p1} - v \right)^2 + 4u \frac{h_p}{2} \tilde{\lambda}_{p1} &= \left(\frac{h_0}{2} \tilde{\lambda}_{p1} - v \right)^2 + 4u \frac{h_0}{2} \tilde{\lambda}_{p1} \\ &\quad + \left(\frac{h_0}{2} \tilde{\lambda}_{p1} + 2u - v \right) \frac{\gamma}{\sqrt{p}} \tilde{\lambda}_{p1} + \frac{\gamma^2}{4p} \tilde{\lambda}_{p1}^2. \end{aligned}$$

Using this expression and the facts that, when $h_p = h_0$,

$$\tilde{\lambda}_{p1} \xrightarrow{a.s.} \frac{(h_0 + 1)(c_1 + h_0)}{h_0(1 + c_1/c_2 + h_0)}, \quad u \rightarrow \frac{1 + c_1/c_2 - c_1}{2}, \quad \text{and } v \rightarrow \frac{1 - c_1}{2},$$

we obtain after some algebra,

$$a_2(h_p) - a_2(h_0) = -\frac{\tilde{\lambda}_{1p}}{4} \left(1 + \frac{\frac{h_0}{2} \tilde{\lambda}_{p1} + 2u - v}{S} \right) \frac{\gamma}{\sqrt{p}} + C^{(2)} \frac{\gamma^2}{p} + o_P\left(\frac{1}{p}\right), \quad (69)$$

where

$$S = \sqrt{\left(\frac{h_0}{2} \tilde{\lambda}_{p1} - v \right)^2 + 4u \frac{h_0}{2} \tilde{\lambda}_{p1}},$$

and

$$C^{(2)} = \frac{c_1 (h_0 + 1)^2 (c_1 + h_0)^2 (c_1 + c_2 - c_1 c_2) (c_2 + c_1 + h_0 c_2)}{2h_0^2 (2c_2 h_0 + c_1^2 + c_2 + c_2 h_0^2 + c_1 + 2h_0 c_1)^3}.$$

For a_3 , we have

$$\begin{aligned} a_3(h_p) - a_3(h_0) = & -\frac{u \tilde{\lambda}_{p1} \left(\frac{h_0}{2} \tilde{\lambda}_{p1} + 2u - v + S \right)}{2 \left(\left(\frac{h_0}{2} \tilde{\lambda}_{p1} - v \right)^2 + 4u \frac{h_0}{2} \tilde{\lambda}_{p1} + \left(\frac{h_0}{2} \tilde{\lambda}_{p1} - v \right) S \right)} \frac{\gamma}{\sqrt{p}} \\ & + C^{(2)} C^{(3)} \frac{\gamma^2}{p} + o_P \left(\frac{1}{p} \right), \end{aligned} \quad (70)$$

where

$$C^{(3)} = \frac{(c_1 + c_2 - c_1 c_2)}{c_2 (c_1 + h_0)} + \frac{(2c_2 h_0 + c_1^2 + c_2 + c_2 h_0^2 + c_1 + 2h_0 c_1) (c_1 + c_2 + c_2 h_0)}{2c_1 c_2 (c_1 + h_0)^2}.$$

Finally, for a_4 , we obtain

$$\begin{aligned} a_4(h_p) - a_4(h_0) = & \frac{(u - v) \tilde{\lambda}_{p1} \left(\frac{h_0}{2} \tilde{\lambda}_{p1} + 2u - v - S \right)}{2 \left(\left(\frac{h_0}{2} \tilde{\lambda}_{p1} - v \right)^2 + 4u \frac{h_0}{2} \tilde{\lambda}_{p1} + \left(-\frac{h_0}{2} \tilde{\lambda}_{p1} - v \right) S \right)} \frac{\gamma}{\sqrt{p}} \\ & - C^{(2)} C^{(4)} \frac{\gamma^2}{p} + o_P \left(\frac{1}{p} \right), \end{aligned} \quad (71)$$

where

$$C^{(4)} = \frac{c_1 + c_2 + c_2 h_0}{c_2 (c_1 + h_0)} + \frac{(c_1 + c_2 - c_1 c_2) (c_1 + c_2 + c_2 h_0^2 + c_1^2 + 2c_1 h_0 + 2c_2 h_0)}{2c_2 (c_1 + h_0)^2 (c_1 + c_2 + c_2 h_0)}.$$

Summing up the γ^2/p terms in the expansions (68-71), we obtain that the γ^2/p term in the expansion of $\sum_{j=1}^4 (a_j(h_p) - a_j(h_0))$, which we will refer as T_2 , equals

$$T_2 = -\frac{1}{4} c_1 \frac{c_1 + c_2 + c_2 h_0^2 - h_0^2 + 2c_2 h_0}{h_0^2 (c_1 + c_2 + c_2 h_0^2 + c_1^2 + 2c_1 h_0 + 2c_2 h_0)} \frac{\gamma^2}{p}. \quad (72)$$

Now let $\Delta = \sqrt{p}(\lambda_{p1} - x_{p1})$, where

$$x_{p1} = \frac{(h_0 + p/n_1)(h_0 + 1)}{h_0 - (h_0 + 1)p/n_2},$$

so that

$$\begin{aligned}\tilde{\lambda}_{p1} &= \frac{\lambda_{p1}}{\frac{n_2}{n_1} + \lambda_{p1}} = \frac{x_{p1} + \Delta/\sqrt{p}}{\frac{n_2}{n_1} + x_{p1} + \Delta/\sqrt{p}} \\ &= \frac{(h_0 + 1)(p/n_1 + h_0)}{h_0(1 + n_2/n_1 + h_0)} + \frac{\Delta}{\sqrt{p}} \frac{c_2 c_1 (c_2 + c_2 h_0 - h_0)^2}{h_0^2 (c_2 + c_1 + c_2 h_0)^2} + o_P\left(\frac{1}{\sqrt{p}}\right),\end{aligned}$$

Our next goal is to expand the weights on γ/\sqrt{p} in expansions (68-71) into power series of Δ/\sqrt{p} up to the linear term only.

For (69), we have

$$-\frac{\tilde{\lambda}_{p1}}{4} \left(1 + \frac{\frac{h_0}{2} \tilde{\lambda}_{p1} + 2u - v}{S} \right) = \tau_0^{(2)} + \left[\tau_{11}^{(2)} + \tau_{12}^{(2)} \right] \frac{\Delta}{\sqrt{p}} + o_P\left(\frac{1}{\sqrt{p}}\right),$$

where

$$\begin{aligned}\tau_{11}^{(2)} &= -\frac{c_1 (c_2 + c_2 h_0 - h_0)^2}{2h_0^2 (c_1 + c_2 + c_2 h_0^2 + c_1^2 + 2c_1 h_0 + 2c_2 h_0)}, \\ \tau_{12}^{(2)} &= \frac{c_1^2 (h_0 + 1) (c_1 + h_0) (c_1 + c_2 - c_1 c_2) (c_2 + c_2 h_0 - h_0)^2}{h_0^2 (c_1 + c_2 + c_2 h_0^2 + c_1^2 + 2c_1 h_0 + 2c_2 h_0)^3},\end{aligned}$$

and $\tau_0^{(2)}$ is a complicated function of h_0, p, n_1 , and n_2 , which we do not report here.

For (70), we have

$$-\frac{u \tilde{\lambda}_{p1} \left(\frac{h_0}{2} \tilde{\lambda}_{p1} + 2u - v + S \right)}{2 \left(\left(\frac{h_0}{2} \tilde{\lambda}_{p1} - v \right)^2 + 4u \frac{h_0}{2} \tilde{\lambda}_{p1} + \left(\frac{h_0}{2} \tilde{\lambda}_{p1} - v \right) S \right)} = \tau_0^{(3)} + \left[\tau_{11}^{(3)} + \tau_{12}^{(3)} \right] \frac{\Delta}{\sqrt{p}} + o_P\left(\frac{1}{\sqrt{p}}\right),$$

where

$$\begin{aligned}\tau_{11}^{(3)} &= \frac{c_1 (c_2 - h_0 + c_2 h_0)^2 (c_1 + c_2 - c_1 c_2) (c_1 + c_2 + c_2 h_0) (c_1 + c_2 + c_2 h_0^2 - c_1^2 + 2c_2 h_0)}{2h_0^2 c_2 (c_1 + c_2 + c_2 h_0^2 + c_1^2 + 2c_1 h_0 + 2c_2 h_0)^3}, \\ \tau_{12}^{(3)} &= -\frac{c_1 (h_0 + 1) (c_2 - h_0 + c_2 h_0)^2 (c_1 + c_2 - c_1 c_2)}{2h_0^2 (c_1 + c_2 + c_2 h_0^2 + c_1^2 + 2c_1 h_0 + 2c_2 h_0)^2},\end{aligned}$$

and $\tau_0^{(3)}$ is a complicated function of h_0, p, n_1 , and n_2 , which we do not report here.

For (71), we have

$$\frac{(u - v) \tilde{\lambda}_{p1} \left(\frac{h_0}{2} \tilde{\lambda}_{p1} + 2u - v - S \right)}{2 \left(\left(\frac{h_0}{2} \tilde{\lambda}_{p1} - v \right)^2 + 4u \frac{h_0}{2} \tilde{\lambda}_{p1} + \left(-\frac{h_0}{2} \tilde{\lambda}_{p1} - v \right) S \right)} = \tau_0^{(4)} + \left[\tau_{11}^{(4)} + \tau_{12}^{(4)} \right] \frac{\Delta}{\sqrt{p}} + o_P\left(\frac{1}{\sqrt{p}}\right),$$

where

$$\tau_{11}^{(4)} = \frac{c_1^3 (c_2 - h_0 + c_2 h_0)^2 (c_1 + c_2 - c_1 c_2) (-c_1 - c_2 + c_2 h_0^2 + c_1^2 + 2c_1 c_2 + 2c_1 c_2 h_0)}{2c_2 h_0^2 (c_1 + c_2 + c_2 h_0) (c_1 + c_2 + c_2 h_0^2 + c_1^2 + 2c_1 h_0 + 2c_2 h_0)^3},$$

$$\tau_{12}^{(4)} = -\frac{c_1^2 (h_0 + 1) (c_2 - h_0 + c_2 h_0)^2 (c_1 + c_2 - c_1 c_2)}{2h_0^2 (c_1 + c_2 + c_2 h_0) (c_1 + c_2 + c_2 h_0^2 + c_1^2 + 2c_1 h_0 + 2c_2 h_0)^2},$$

and $\tau_0^{(4)}$ is a complicated function of h_0, p, n_1 , and n_2 , which we do not report here.

We have verified, using Maple symbolic algebra software, that

$$\tau^{(2)} + \tau^{(3)} + \tau^{(4)} = -\frac{1}{2} \frac{1 + h_0}{h_0},$$

which is exactly the negative of the term on γ/\sqrt{p} in (68). Hence, the term on γ/\sqrt{p} in the expansion of $\sum_{j=1}^4 (a_j(h_p) - a_j(h_0))$ is zero. Further, we have verified that

$$\sum_{j=2}^4 (\tau_{11}^{(j)} + \tau_{12}^{(j)}) = -\frac{1}{2} c_1 \frac{(c_2 - h_0 + c_2 h_0)^2}{h_0^2 (c_1 + c_2 + c_2 h_0^2 + c_1^2 + 2c_1 h_0 + 2c_2 h_0)}.$$

This equality, together with (67) and (72) imply that

$$\begin{aligned} \ln L_{p1}(\gamma, \Lambda) &\stackrel{P}{\sim} \frac{1}{2} \frac{(c_2 - h_0 + c_2 h_0)^2}{h_0^2 (c_1 + c_2 + c_2 h_0^2 + c_1^2 + 2c_1 h_0 + 2c_2 h_0)} \gamma \Delta \\ &+ \frac{1}{4} \frac{c_1 + c_2 + c_2 h_0^2 - h_0^2 + 2c_2 h_0}{h_0^2 (c_1 + c_2 + c_2 h_0^2 + c_1^2 + 2c_1 h_0 + 2c_2 h_0)} \gamma^2. \end{aligned} \quad (73)$$

Consider a different local parameter

$$\theta_2 = \gamma/\omega_2(h_0),$$

where

$$\omega_2(h_0) = \frac{2h_0^2 (c_1 + c_2 + c_2 h_0^2 + c_1^2 + 2c_1 h_0 + 2c_2 h_0)}{(h_0 - c_2(1 + h_0))^2}.$$

Asymptotic approximation (73) implies the following lemma.

Lemma 20 *Under the null hypothesis that $h = h_0$, uniformly in θ_2 from any compact subset of \mathbb{R} ,*

$$\ln L_{p2}(\gamma, \Lambda) = \theta_2 \sqrt{p} (\lambda_{p1} - x_{p1}) - \frac{1}{2} \theta_2^2 \tau_2^2(h_0) + o_P(1)$$

where

$$x_{p1} = \frac{(h_0 + p/n_1)(h_0 + 1)}{h_0 - (h_0 + 1)p/n_2}, \text{ and}$$

$$\tau_2^2(h_0) = \frac{2h_0^2 \left(h_0^2 - c_2(1 + h_0)^2 - c_1 \right) \left((c_1 + c_2)(1 + h_0)^2 - c_1(h_0^2 - c_1) \right)}{(c_2 - h_0 + c_2 h_0)^4}.$$

Similarly to the case of Setting 1, Lemma 20 together with the asymptotic normality of $\sqrt{p}(\lambda_{p1} - x_{p1})$ established in Proposition 11 imply, via Le Cam's First Lemma (see van der Vaart (1998), p.88), that the sequences of the probability measures $\{\mathbb{P}_{h_0,p}\}$ and $\{\mathbb{P}_{h_0+\gamma/\sqrt{p},p}\}$ describing the joint distribution of the eigenvalues of \mathbf{F} under the null $H_0 : h_p = h_0$ and under the local alternative $H_1 : h_p = h_0 + \gamma/\sqrt{p}$ are mutually contiguous. Moreover, the experiments $(\mathbb{P}_{h_0+\theta_2\omega_2(h_0)/\sqrt{p},p} : \theta_2 \in \mathbb{R})$ converge to the Gaussian shift experiment $(N(\theta_2, \tau_2^2(h_0)) : \theta_2 \in \mathbb{R})$. In particular, these experiments are LAN.

7 Conclusion

In this paper, we establish the Local Asymptotic Normality of the experiments of observing the eigenvalues of the F-ratio $\mathbf{F} \equiv (B/n_2)^{-1} A/n_A$ of two large-dimensional Wishart matrices. The experiments are parameterized by the value of a single spike that describes the “ratio” of the covariance parameters of A and B , or, in the case of equal covariance parameters, the non-centrality parameter of A . We find that the asymptotic behavior of the log ratio of the joint density of the eigenvalues of \mathbf{F} , which corresponds to a super-critical spike, to their joint density under a local deviation from this value depends only on the largest eigenvalue λ_{p1} . This implies, in particular, that the best statistical inference about a super-critical spike in the local asymptotic regime is based on the largest eigenvalue only.

As a by-product of our analysis, in a multi-spike setting, we establish the joint asymptotic normality of a few of the largest eigenvalues of \mathbf{F} that correspond to the super-critical spikes. We derive an explicit formulas for the almost sure limits of these eigenvalues, and for the asymptotic variances of their fluctuations around these limits.

8 Acknowledgements

This work was supported in part by NIH grant 5R01 EB 001988 (PD, IMJ), the Simons Foundation Math + X program (PD), NSF grant DMS 1407813 (IMJ), and the the J.M. Keynes Fellowships Fund, University of Cambridge (AO).

9 Appendix

9.1 Proof of Lemma 8

We will need the following two lemmas.

Lemma 21 (McLeish 1974) *Let $\{X_{pj}, \mathcal{G}_{pj}, j = 1, \dots, p\}$ be a martingale difference array on the probability triple (Ω, \mathcal{G}, P) . If the following conditions are satisfied: a) Lindeberg's condition: for all $\varepsilon > 0$, $\sum_j \int_{|X_{pj}| > \varepsilon} X_{pj}^2 dP \rightarrow 0$ as $p \rightarrow \infty$; b) $\sum_j X_{pj}^2 \xrightarrow{P} 1$, then $\sum_j X_{pj} \xrightarrow{d} N(0, 1)$.*

Proof: This is a consequence of Theorem (2.3) of McLeish (1974). Two conditions of the theorem: i) $\max_{j \leq p} |X_{pj}|$ is uniformly bounded in L_2 norm, and ii) $\max_{j \leq p} |X_{pj}| \xrightarrow{P} 0$, are replaced here by the Lindeberg condition. \square

Lemma 22 (Hall and Heyde) *Let $\{X_{pj}, \mathcal{G}_{pj}, j = 1, \dots, p\}$ be a martingale difference array, and define $V_{pJ}^2 = \sum_{j=1}^J E(X_{pj}^2 | \mathcal{G}_{p,j-1})$ and $U_{pJ}^2 = \sum_{j=1}^J X_{pj}^2$ for $J = 1, \dots, p$. Suppose that the conditional variances V_{pp}^2 are tight, that is $\sup_p P(V_{pp}^2 > \varepsilon) \rightarrow 0$ as $\varepsilon \rightarrow \infty$, and that the conditional Lindeberg condition holds, that is, for all $\varepsilon > 0$, $\sum_j E[X_{pj}^2 \mathbf{1}\{|X_{pj}| > \varepsilon\} | \mathcal{G}_{p,j-1}] \xrightarrow{P} 0$. Then $\max_J |U_{pJ}^2 - V_{pJ}^2| \xrightarrow{P} 0$.*

Proof: This is a shortened version of Theorem 2.23 in Hall and Heyde (1980). \square

Let $f_q(\lambda)$, $q = 1, \dots, Q$, be such that $f_q(\lambda) = g_q(\lambda)$ for $\lambda \in [l_i, L_i]$ and $f_q(\lambda) = 0$ otherwise. Consider random variables

$$X_{pj} = \frac{1}{\sqrt{p}} \sum_{(q,s,t) \in \Theta} \gamma_{qst} f_q \left(\lambda_{pj}^{(i)} \right) (\zeta_{js} \zeta_{jt} - \delta_{st}),$$

where γ_{qst} are some constants. Let \mathcal{G}_{pJ} be the σ -algebra generated by $\lambda_{p1}^{(i)}, \dots, \lambda_{pp}^{(i)}$ and ζ_{js} with $j = 1, \dots, J$; $s = 1, \dots, m$. Clearly, $\{X_{pj}, \mathcal{G}_{pj}, j = 1, \dots, p\}$ form a martingale difference array. Let K be the number of different triples $(q, s, t) \in \Theta$. Consider an arbitrary order in Θ . In Hölder's inequality

$$\sum_{a=1}^K y_a z_a \leq \left(\sum_{a=1}^K (y_a)^b \right)^{1/b} \left(\sum_{a=1}^K (z_a)^c \right)^{1/c},$$

which holds for $y_a > 0$, $z_a > 0$, $b > 1$, $c > 1$, and $1/b + 1/c = 1$, take

$$y_a = \left| \frac{1}{\sqrt{p}} \gamma_{qst} f_q \left(\lambda_{pj}^{(i)} \right) (\zeta_{js} \zeta_{jt} - \delta_{st}) \right|,$$

where (q, s, t) is the a -th triple in Θ , $z_a = 1$, and $b = 2 + \delta$ for some $\delta > 0$. Then, the inequality implies that

$$|X_{pj}|^{2+\delta} \leq K^{1+\delta} R_i^{2+\delta} \sum_{(q,s,t) \in \Theta} \left| \frac{1}{\sqrt{p}} \gamma_{qst} (\zeta_{js} \zeta_{jt} - \delta_{st}) \right|^{2+\delta}, \quad (74)$$

where

$$R_i = \max_{q=1, \dots, Q} \sup_{\lambda \in [l_i, L_i]} |g_q(\lambda)|.$$

Since ζ_{js} are i.i.d. $N(0, 1)$, (74) implies that $\sum_{j=1}^p E |X_{pj}|^{2+\delta} \rightarrow 0$ as $p \rightarrow \infty$, which means that the Lyapunov condition holds for X_{pj} . As is well known, Lyapunov's condition implies Lindeberg's condition. Hence, condition a) of Lemma 21 is satisfied for X_{pj} .

Let us consider $\sum_{j=1}^p X_{pj}^2$. Since the convergence in mean implies the convergence in probability, the conditional Lindeberg condition is satisfied for X_{pj} because the unconditional Lindeberg condition is satisfied as checked above. Further, in notations of Lemma 22, it is easy to see that

$$V_{pp}^2 = \sum_{q, q_1} \left[\left(\sum_{1 \leq s \leq t \leq m} \gamma_{qst} \gamma_{q_1 st} (1 + \delta_{st}) \right) \frac{1}{p} \sum_{j=1}^p f_q \left(\lambda_{pj}^{(i)} \right) f_{q_1} \left(\lambda_{pj}^{(i)} \right) \right].$$

The convergence of the empirical distribution of $\lambda_{p1}^{(i)}, \dots, \lambda_{pp}^{(i)}$ to G_{x_i} and the equality of g_q and f_q on the support of G_{x_i} implies that

$$V_{pp}^2 \xrightarrow{P} \Sigma \equiv \sum_{q, q_1} \left[\left(\sum_{1 \leq s \leq t \leq m} \gamma_{qst} \gamma_{q_1 st} (1 + \delta_{st}) \right) \int g_q(\lambda) g_{q_1}(\lambda) dG_{x_i} \right].$$

In particular, V_{pp}^2 is tight and Lemma 22 applies. Therefore, $\sum_{j=1}^p X_{pj}^2$ converges to the same limit as V_{pp}^2 . Thus, by Lemma 21, we get $\sum_{j=1}^p X_{pj} \xrightarrow{d} N(0, \Sigma)$.

Finally, let

$$Y_{pj} = \frac{1}{\sqrt{p}} \sum_{(q,s,t) \in \Theta} \gamma_{qst} g_q \left(\lambda_{pj}^{(i)} \right) (\zeta_{js} \zeta_{jt} - \delta_{st}).$$

Since

$$\Pr \left(\sum_{j=1}^p X_{pj} \neq \sum_{j=1}^p Y_{pj} \right) \rightarrow 0$$

as $p \rightarrow \infty$, we have $\sum_{j=1}^p Y_{pj} \xrightarrow{d} N(0, \Sigma)$. Lemma 8 follows from this convergence via the Cramer-Wold device. \square

9.2 Derivation of (29), (30), and (31)

Expression (29) immediately follows from (15). Next, differentiating identity (13) with respect to z , we obtain

$$1 + \frac{c_1 m'_x(z)}{(1 + c_1 m_x(z))^2} = \frac{m'_x(z)}{m_x^2(z)} + \frac{-x^2 c_2 m'_x(z)}{(1 - c_2 x m_x(z))^2}.$$

Setting $z = 0$ and $x = x_i$, and using the fact that

$$m_{x_i}(0) = -(h_i + c_1)^{-1}, \quad (75)$$

which follows from (15), we obtain

$$1 + \frac{c_1 m'_{x_i}(0)}{(1 - c_1 (h_i + c_1)^{-1})^2} = \frac{m'_{x_i}(0)}{(h_i + c_1)^{-2}} + \frac{-x_i^2 c_2 m'_{x_i}(0)}{(1 + c_2 x_i (h_i + c_1)^{-1})^2}.$$

Using the definition (17) of x_i , we obtain

$$1 + \frac{c_1 m'_{x_i}(0)}{(1 - c_1 (h_i + c_1)^{-1})^2} = \frac{m'_{x_i}(0)}{(h_i + c_1)^{-2}} - \frac{(h_i + c_1)^2 (h_i + 1)^2 c_2 m'_{x_0}(0)}{h_i^2},$$

which implies (30). Finally, differentiating identity (13) with respect to x , we obtain

$$\begin{aligned} \frac{c_1 dm_x(z)/dx}{(1 + c_1 m_x(z))^2} &= \frac{dm_x(z)/dx}{(m_x(z))^2} \\ &+ \frac{-1 + c_2 x m_x(z) - x(c_2 m_x(z) + c_2 x dm_x(z)/dx)}{(1 - c_2 x m_x(z))^2}. \end{aligned}$$

Setting $z = 0$ and $x = x_i$, we obtain

$$\frac{c_1 dm_{x_i}(0)/dx}{(1 + c_1 m_{x_i}(0))^2} = \frac{dm_{x_i}(0)/dx}{(m_{x_i}(0))^2} + \frac{-1 - c_2 x_i^2 dm_{x_i}(0)/dx}{(1 - c_2 x_i m_{x_i}(0))^2}.$$

This equality, the definition (17) of x_i , and equation (75) imply (31).

References

- [1] Bai, Z.D. and Silverstein, J.W. (1998) “No Eigenvalues Outside the Support of the Limiting Spectral Distribution of Large-Dimensional Sample Covariance Matrices,” The Annals of Probability, Vol. 26, pp. 316-345.

- [2] Bai, Z.D. and J. Yao (2008) “Central limit theorems for eigenvalues in a spiked population model,” *Annales de l’Institut Henri Poincaré - Probabilités et Statistiques* 44, 447–474.
- [3] Baik, J., G. Ben Arous and S. Péché (2005) “Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices,” *Annals of Probability* 33, 1643–1697.
- [4] Bao, Z., J. Hu, G. Pan, and W. Zhou (2014b) “Canonical correlation coefficients of high-dimensional normal vectors: finite rank case,” *arXiv* 1407.7194
- [5] Bao, Z., G. Pan, and W. Zhou (2014a) “Universality for the largest eigenvalue of sample covariance matrices with general population,” *arXiv*: 1304.5690v6.
- [6] Benaych-Georges, F., A. Guionnet, and M. Maida (2011) “Fluctuations of the Extreme Eigenvalues of Finite Rank Deformations of Random Matrices,” *Electronic Journal of Probability* 16, 1621-1662.
- [7] Bilodeau, M. and M. S. Srivastava (1992) “Estimation of the Eigenvalues of $\Sigma_1 \Sigma_2^{-1}$,” *Journal of Multivariate Analysis* 41, 1-13.
- [8] Bloemendal, A. and B. Virág (2013) “Limits of spiked random matrices I,” *Probability Theory and Related Fields* 156, 795-825.
- [9] Bloemendal, A. and B. Virág (2011) “Limits of spiked random matrices II,” *arXiv*: 1109.3704v1.
- [10] Dharmawansa, P. and I. M. Johnstone (2014) “Joint density of eigenvalues in spiked multivariate models,” *Stat* 3, no. 1, 240–249.
- [11] Féral, D. and S. Péché (2009) “The largest eigenvalues of sample covariance matrices for a spiked population: diagonal case,” *Journal of Mathematical Physics* 50, 073302.
- [12] Hall, P., and C.C. Heyde (1980) *Martingale limit theory and its application*, New York: Academic Press.
- [13] J. F. Hayes, J. F. and W. G. Hill (1981) “Modification of Estimates of Parameters in the Construction of Genetic Selection Indices (‘Bending’),” *Biometrics* Vol. 37, No. 3, pp. 483-493
- [14] Horn, R. A., and C. R. Johnson (1985) *Matrix Analysis*, Cambridge University Press.
- [15] James, A. T. (1964) “Distributions of matrix variates and latent roots derived from normal samples”, *Annals of Mathematical Statistics* 35, 475-501.

- [16] Johnstone, I. M. and A. Onatski (2014) “Likelihood ratio analysis in the sub-critical regime. Case 4,” manuscript Case4version7.pdf, in preparation.
- [17] Kato, T. (1980) *Perturbation Theory for Linear Operators*, Springer-Verlag. Berlin, Heidelberg, New York.
- [18] Khatrı, C. G. (1967) “Some distributional problems associated with the characteristics roots of $S_1 S_2^{-1}$,” *Ann. Math. Stat.*, vol. 38, no. 3, pp. 944–948.
- [19] Kritzman, M. and Y. Li (2010) “Skulls, Financial Turbulence, and Risk Management,” *Financial Analysts Journal* 66, 30-41.
- [20] McLeish, D.L. (1974) “Dependent Central Limit Theorems and Invariance Principles”, *Annals of Probability*, Vol. 2, No. 4, p.620-628.
- [21] Mo, M.Y. (2012) “The rank 1 real Wishart spiked model,” *Communications on Pure and Applied Mathematics* 65, 1528–1638.
- [22] Muirhead, R.J. (1982) *Aspects of Multivariate Statistical Theory*. John Wiley & Son, Hoboken, New Jersey.
- [23] Muirhead, R.J. and T. Verathaworn (1985) “On estimating the latent roots of $\Sigma_1 \Sigma_2^{-1}$,” in Krishnaiah, P.R. (ed.) *Multivariate Analysis - VI*, Elsevier Science Publishers B.V., 431-447.
- [24] Nadakuditi, R.R. and J. W. Silverstein (2010) “Fundamental Limit of Sample Generalized Eigenvalue Based Detection of Signals in Noise Using Relatively Few Signal-Bearing and Noise-Only Samples,” *IEEE Journal of Selected Topics in Signal Processing* 4 (3), 468-480.
- [25] Olver, F.W.J. (1997) *Asymptotics and Special Functions*, A K Peters, Natick, Massachusetts.
- [26] Onatski, A. (2007) “Asymptotics of the principal components estimator of large factor models with weak factors and i.i.d. Gaussian noise,” manuscript, University of Cambridge. Available at <http://www.econ.cam.ac.uk/people/faculty/ao319/pubs/inference45a.pdf>.
- [27] Onatski, A. (2012) “Asymptotics of the principal components estimator of large factor models with weakly influential factors,” *Journal of Econometrics* 168, 244-258.
- [28] Onatski, A. (2012a) “Supplementary Appendix to Onatski (2012)”, available at <http://www.econ.cam.ac.uk> in the section People, Dr Alexei Onatski, Link To My Papers

- [29] Onatski, A., Moreira, M.J., and M. Hallin (2013) “Asymptotic power of sphericity tests for high-dimensional data,” *Annals of Statistics* 41, 1204-1231.
- [30] Onatski, A., Moreira, M.J., and M. Hallin (2014) “Signal Detection in High Dimension: The Multispiked Case,” *Annals of Statistics* 42, 225-254.
- [31] Passemier, D., M. R. McKay, and Y. Chen (2014) “Hypergeometric Functions of Matrix Arguments and Linear Statistics of Multi-Spiked Hermitian matrix Models,” arXiv: 1406.0791v2.
- [32] Paul, D. (2007) “Asymptotics of the leading sample eigenvalues for a spiked covariance model,” *Statistica Sinica* 17, 1617–1642.
- [33] Rudin, W. (1987) *Real and Complex Analysis*, 3-rd edition. McGraw-Hill series in higher mathematics.
- [34] Sheena, Y., A.K. Gupta, and Y. Fujikoshi (2004) “Estimation of the eigenvalues of noncentrality parameter in matrix variate noncentral beta distribution,” *Annals of the Institute of Statistical Mathematics* 56, 101-125.
- [35] Silverstein, J.W. (1995) “Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices,” *Journal of Multivariate Analysis* 5, 331-339.
- [36] Silverstein, J.W. and Bai, Z.D. (1995) “On the empirical distribution of eigenvalues of large dimensional random matrices,” *Journal of Multivariate Analysis* 54, 175-192.
- [37] Silverstein, J.W. and S. Choi (1995) “Analysis of the Limiting Spectral Distribution of Large Dimensional Random Matrices,” *Journal of Multivariate Analysis*, 54, 295-309
- [38] Titchmarsh, E. C. (1960) *The Theory of Functions*, 2nd ed. Oxford, England. Oxford University Press.
- [39] van der Vaart, A.W. (1998) *Asymptotic Statistics*. Cambridge University Press.
- [40] Wachter, K. (1980) “The limiting empirical measure of multiple discriminant ratios,” *The Annals of Statistics* 8, 937-957.